AFRL-RV-PS-
TP-2014-0006

PIPELINE PROCESSING WITH AN ITERATIVE, CONTEXT-BASED DETECTION MODEL
Annual Report

T. Kværna, et al.

NORSAR
PO Box 53
N-2027 Kjeller, Norway

19 April 2014

Technical Report

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION IS UNLIMITED.

AIR FORCE RESEARCH LABORATORY
Space Vehicles Directorate
3550 Aberdeen Ave SE
AIR FORCE MATERIEL COMMAND
KIRTLAND AIR FORCE BASE, NM 87117-5776

DTIC COPY

NOTICE AND SIGNATURE PAGE

Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications, or other data does not license the holder or any other person or corporation; or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

This report was cleared for public release by the 377 ABW Public Affairs Office and is available to the general public, including foreign nationals. Copies may be obtained from the Defense Technical Information Center (DTIC) (http://www.dtic.mil).

AFRL-RV-PS-TP-2014-0006 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

//SIGNED//	//SIGNED//
Robert Raistrick	Glenn M. Vaughan, Colonel, USAF
Project Manager, AFRL/RVBYE	Chief, Battlespace Environment Division

This report is published in the interest of scientific and technical information exchange, and its publication does not constitute the Government's approval or disapproval of its ideas or findings.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
19-04-2014	Technical Report	19 Mar 2013 to 18 Mar 2014

4. TITLE AND SUBTITLE
PIPELINE PROCESSING WITH AN ITERATIVE, CONTEXT-BASED DETECTION MODEL
Annual Report

5a. CONTRACT NUMBER
FA9453-13-C-0270

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
62601F

6. AUTHOR(S)
T. Kværna, D. A. Dodge[1], D. B. Harris[2], and S. J. Gibbons

5d. PROJECT NUMBER
1010

5e. TASK NUMBER
PPM00018850

5f. WORK UNIT NUMBER
EF122183

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
NORSAR
PO Box 53
N-2027 Kjeller, Norway

[1]Lawrence Livermore National Laboratory
7000 East Avenue
Livermore, CA 94550

[2] Deschutes Signal Processing, LLC
81211 E. Wapinitia Road
Maupin, Oregon 97037

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Research Laboratory
Space Vehicles Directorate
3550 Aberdeen Avenue SE
Kirtland AFB, NM 87117-5776

10. SPONSOR/MONITOR'S ACRONYM(S)
AFRL/RVBYE

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
AFRL-RV-PS-TP-2014-0006

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited. (377ABW-2014-1006 dtd 09 Dec 2014)

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Under existing detection pipelines, seismic event hypotheses are formed from a parametric description of the waveform data obtained from a single pass over the incoming data stream. The full potential of signal processing algorithms is not being exploited due to simplistic assumptions made about the background against which signals are being detected. A vast improvement in the available computational resources allows the possibility of more sensitive and more robust context-based detection pipelines which glean progressively more information from multiple passes over the data. In the first year of this two year contract we have designed and implemented several extensions to an existing prototype detection framework to demonstrate the feasibility of improving performance from a systematic reprocessing of the raw data. The new components are: signal cancellation for stripping the incoming data stream of repeating and irrelevant signals prior to running primary detectors, adaptive beamforming and matched field processing for suppressing background signals and aftershock sequences, and the testing of event hypotheses by evaluating detection probabilities for both detecting and non-detecting stations, followed by optimized beamforming.

15. SUBJECT TERMS
aftershock sequences, repeating explosions, detection framework, pattern detectors, correlation detectors, subspace detectors, matched field detectors, nuclear explosion monitoring

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Robert Raistrick
Unclassified	Unclassified	Unclassified	Unlimited	52	**19b. TELEPHONE NUMBER** *(include area code)*

This page is intentionally left blank.

Approved for public release; distribution is unlimited.

Table of Contents

Approved for public release; distribution is unlimited.

List of Figures

Approved for public release; distribution is unlimited.

Approved for public release; distribution is unlimited.

This page is intentionally left blank.

Approved for public release; distribution is unlimited.

1 SUMMARY

Under existing detection pipelines, seismic event hypotheses are formed from a parametric description of the waveform data obtained from a single pass over the incoming data stream. The full potential of signal processing algorithms is not being exploited due to simplistic assumptions made about the background against which signals are being detected. A vast improvement in the available computational resources allows the possibility of more sensitive and more robust context-based detection pipelines which glean progressively more information from multiple passes over the data. In the first year of this two year contract we have designed and implemented several extensions to an existing prototype detection framework to demonstrate the feasibility of improving performance from a systematic reprocessing of the raw data. The new components are: signal cancellation for stripping the incoming data stream of repeating and irrelevant signals prior to running primary detectors, adaptive beamforming and matched field processing for suppressing background signals and aftershock sequences, and the testing of event hypotheses by evaluating detection probabilities for both detecting and non-detecting stations, followed by optimized beamforming.

2 INTRODUCTION

The traditional processing flow in monitoring pipelines consists of a transformation of continuous waveform data into a stream of parametric information describing discrete detected arrivals. Critical event-building functions (association, location and identification) are performed subsequently on the parametric data. This processing architecture is a legacy of computing resources available during the 1980s and 1990s, for which very intensive signal processing operations were computationally prohibitive. Expensive signal processing operations were minimized in such systems by performing only a single pass over the waveform data.

One consequence of this architecture is that only very simple assumptions are possible about ambient background conditions in the selection and development of signal processing algorithms. The key assumptions – clearly false in important and not-infrequent circumstances – are that events occur in isolation and that the background consists of white noise, usually uncorrelated and uniform in power among array sensors. Pipelines do have procedures to disentangle overlapped and interleaved seismic arrivals, but these usually are confined to association algorithms operating on the parametric representation of detected arrivals. Except to allow analysts to verify whether events have been built properly, and build new ones as necessary, monitoring systems generally do not revisit waveform data once it has been reduced

Approved for public release; distribution is unlimited.

to its parametric description.

This single-pass architecture imposes substantial limitations on the signal processing algorithms that are applied for detection and phase characterization. Among these is the inability to adapt algorithms to suppress or cancel interfering signals with known characteristics. Future pipelines could allow waveform data to be revisited by detection and characterization algorithms that exploit some understanding or model of the context in which they operate. We present two examples suggesting that significant performance improvements are possible with context-driven signal processing. The work plan includes an examination of how widely these innovations might enhance monitoring operations. However, full consideration or modeling of an advanced pipeline architecture is beyond the scope of this contract. Consequently, we limit our investigation to the implications of an iterative, context-driven processing approach for beamforming and related detection algorithms.

In contemplating the signal processing front-end of future pipelines, we assume the overall systems would have characteristics or functions not present in current implementations. For example, we imagine that central data collection systems would have autonomous components to characterize local sources of interfering signals near each station and remove or suppress those signals from the data stream before it is passed to the main pipeline for detection and subsequent processing. Such a subsystem would generalize existing masking functions that handle dropouts and data spikes to handle interfering, propagating transients from local nuisance sources. Examples of such sources abound: glaciers near SPITS producing icequakes, cultural sources near KSRS producing lone Rg phases. At times, these stations can be plagued with thousands of nuisance transients. We give an example of icequake suppression by correlation detection and waveform cancellation. Under the work plan, we will build and test a prototype detection and cancellation component.

Our second assumption about future pipelines is that they would maintain constantly-refined models (i.e. world views) of seismic activity around the globe, and use those models to adapt the signal processing functions of the pipeline to optimize detection and characterization functions. In such architectures, model state might be determined initially by pipeline processes not very different from those used currently. However, waveform data could be extensively reprocessed with algorithms conditioned on hypotheses about the current state of seismic activity (e.g. large aftershock sequence in progress, mid-day mining explosions anticipated, etc.). A mature picture of activity may emerge after several iterations of model-driven signal processing on the continuous data stream. It is not that this picture of activity is of direct monitoring interest, but rather that it may allow more effective screening of interference in the continuous data stream.

Approved for public release; distribution is unlimited.

Context-driven signal processing operations might include beamforming algorithms that search for events of interest among strong transient interference. Current beamforming algorithms are designed for very benign, white, uncorrelated noise background conditions. They are not well suited to detection among events in an aftershock sequence, for example. However, it should be possible to configure beamforming and related (for example, matched field processing) operations in second passes over data to make use of knowledge of interfering events gleaned from an initial pass. We give an example of such an algorithm: a data-adaptive, empirical matched field processor that substantially suppresses aftershock signals while remaining sensitive to signals from a known test site. The capability of this algorithm to suppress interference is significantly beyond the ability of a conventional beamforming algorithm because it is specifically designed to reject events with known characteristics. Our work plan includes a test of this and related algorithms on a substantial fraction of a network with data recorded during an aftershock sequence.

The model-building supervisory function of future pipeline architectures also could instigate signal processing operations to support or refute hypotheses being formed or tested. For example, having formed a marginally-supported event hypothesis with a minimum of phase detections, the supervisor could direct special beams to search for phases predicted to exist at stations without detections. Current architectures have some capability to perform this function. For example, the IDC system performs associations, forms and locates events based on early-arriving seismic and hydroacoustic data (forming the SEL1 bulletin). The system automatically requests data from auxiliary stations based upon event hypotheses reported in that bulletin, and performs detections and measurements on those new data. However, this function can be substantially enhanced to direct re-examination of all waveform data, both to search for missing phases as noted and to confirm detections in the context of an hypothesized event (with location and magnitude estimates). Our work plan includes a test of context-driven detection operations using information on station sensitivity amassed by NORSAR for the IMS network.

The general objective of this project is to examine whether a processing approach that allows waveform data to be revisited repeatedly can substantially improve detection performance in a monitoring pipeline. The specific objectives are to determine whether:

1) information about transients from local sources surrounding an array can be systematically discovered by an autonomous system and exploited to reduce interference with beamforming operations,
2) an advanced pipeline could configure adaptive beams to reject aftershocks more effectively than conventional recipe beams,

Approved for public release; distribution is unlimited.

3) catalog events could be more effectively built by reprocessing data with beams specifically designed to search for phases predicted to exist under event hypotheses but which were not detected on a first pass over the data with conventional recipe beams,
4) false associations can be discovered with specialized beams designed to check phase detections made with conventional recipe beams.

The specialized beams or detection algorithms we have in mind for reprocessing the data in pursuit of objectives 3 and 4 are likely to be adaptive beams configured to search for specific phases in correlated background noise or among waveforms from competing events. Since they would be applied in a second or third pass over the data, they would be designed to detect phases with specified characteristics while rejecting signals already detected and characterized in a previous pass over the data.

Fig. 1 provides an overview of the pipeline architecture whereby the new components to be evaluated here are shaded. Sections (3.1) and (3.2) describe in more detail the objectives (1) and (2) above respectively: the progress achieved in the first year of the contract for these two objectives is detailed in sections (4.1) and (4.2). Section (3.3) describes in greater detail the task of improving the evaluation of event hypotheses (encompassing objectives (3) and (4) above) and the progress achieved in this task in the first year of the contract is described in section (4.3).

Approved for public release; distribution is unlimited.

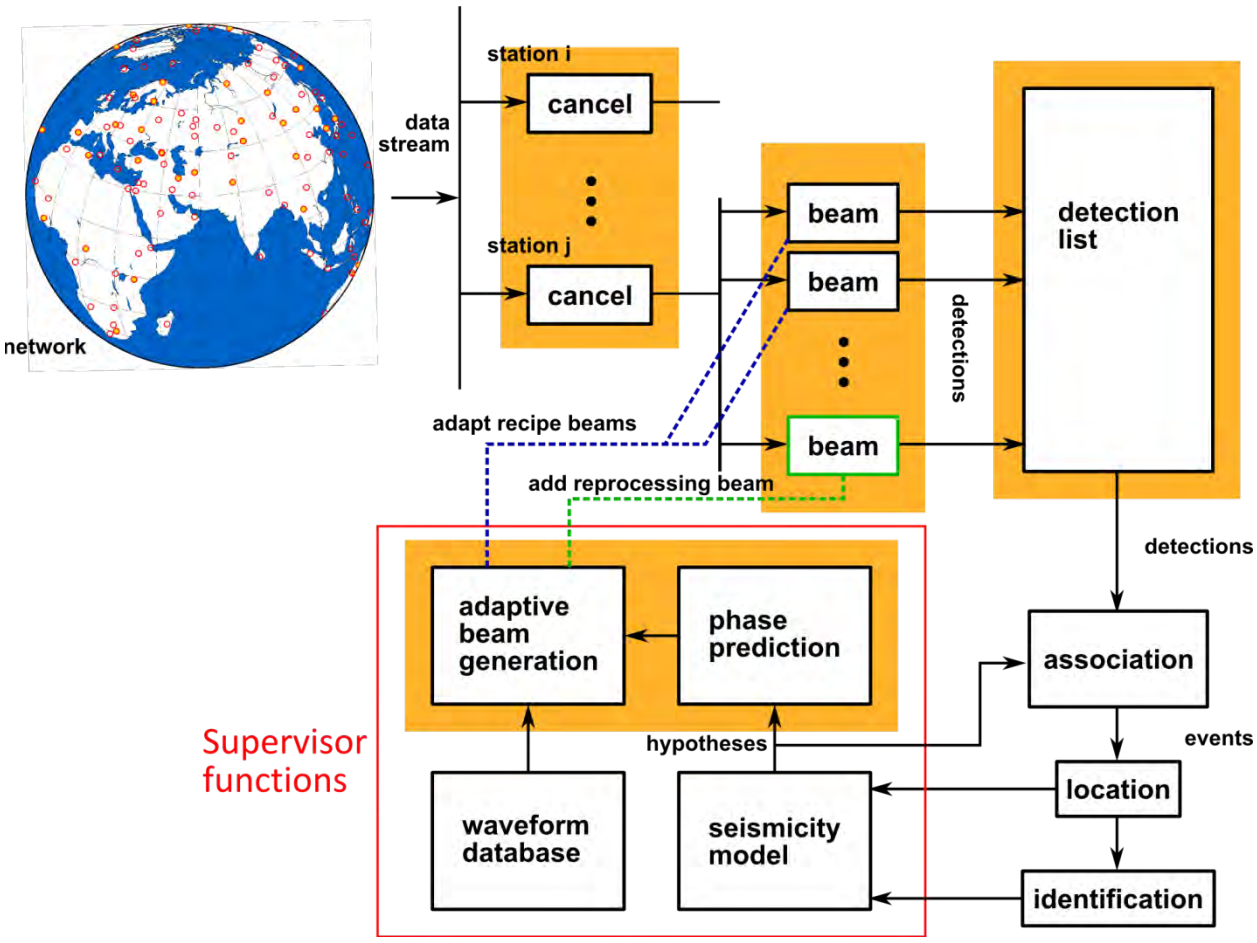

Fig. 1 High level sketch of possible future pipeline functions that would drive context-based signal processing algorithms. *Functions highlighted with an orange backdrop are ones that are being prototyped as part of this project.*

Approved for public release; distribution is unlimited.

3 TECHNICAL APPROACH

The introduction outlines a number of specific procedures that a context-based detection model can incorporate to produce a higher fidelity description of relevant sources of seismicity on local, regional, and global scales. Here we deal with three of these procedures: signal cancellation for problematic transients, adaptive beamforming with matched field detection for aftershock suppression, and the testing of seismic event hypotheses through the reprocessing of raw data.

3.1 Cancellation of local transient interference; a first layer in a hierarchical pipeline

The advent of inexpensive large-scale computing in the next few years provides an opportunity for unparalleled innovation in pipeline architectures. This project is examining the possibility that cheap parallel machines will allow multiple passes over the data by context-driven detection algorithms. The multiple passes may be organized as a hierarchy of pipelines, with the lowest level in the hierarchy consisting of large numbers of locally-focused pipelines (one per station). These might generalize the current pipeline function that pre-processes data streams to identify (mark) glitches and dropouts to make subsequent processing more robust. The kind of generalization we have in mind is one requiring inferences about sources local to a station that produce repeating transient signals of no monitoring interest. Such transients might not be flagged by traditional glitch- or dropout-detecting routines as they are truly propagating signals. In a traditional pipeline, they may be passed on to produce detections and create problems with association algorithms. To the extent that a local pipeline processor can autonomously detect, flag, or even remove transients from the stream, the overall system may be less susceptible to large numbers of undesirable transients.

For nuclear explosion monitoring, the problem of frequent interfering transients from very local sources is of general concern. For example, we are aware that during certain time intervals, numerous short-duration, high SNR, signals from local sources are observed at the KSRS array in South Korea. Other examples are found at arrays located in environments with large diurnal temperature variations (e.g. SPITS, ARCES, FINES, ILAR, NORSAR and YKA). As part of the freeze-thaw cycle, seismic emissions are generated in the upper soil layer close to or even within the arrays. Stations located close to glaciers are prone to recording icequakes, sometimes in the thousands.

Approved for public release; distribution is unlimited.

We are investigating an autonomous sub-pipeline that identifies repeating sources influencing the data stream from a single station (which may be a multichannel stream from an array), characterizes the (typically short) transients these sources produce, then detects such transient occurrences, estimates their waveforms and subtracts them from the stream. The general term we are using for this strategy is cancellation. The "cleaned" stream resulting from cancellation would be passed on to more traditional detectors and aggregated with other local streams into higher levels of the pipeline hierarchy. In keeping with a conservative stance on data modification by autonomous components, the system also keeps a complete trail of metadata describing the processing applied to the stream for use by higher level components of the pipeline.

The principal idea behind cancellation is that many interfering local events are highly repetitive, with waveforms falling into a relatively few patterns. The occurrence and timing of these events can be discovered with correlation detectors. If the waveforms of the new events are high-fidelity clones of a pattern, then it is possible to align the pattern waveform to the occurrence of the new events, scale the template appropriately and subtract it at the right time from the continuous stream to eliminate the interfering signals. Direct implementation of this simple idea works to a degree, but rarely leads to a sufficiently small residual. Imperfect cancellation with a single template waveform can be due to small sampling offsets between the template waveform and the waveform to be cancelled.

A more sophisticated approach using subspace representations [Harris, 1991] for the repeating waveform patterns gives more satisfactory results. In this approach, an ensemble of instances of the pattern event is detected with a correlation detector, and an orthonormal basis for the ensemble waveforms is constructed. A waveform from an undesired interfering event is approximated by a least-squares linear combination of the basis waveforms. This best approximation is subtracted from the continuous stream at the appropriate time to cancel the undesired signal. This approach succeeds where straightforward attempts to cancel using a single template waveform underperform, because the ensemble of waveforms captures timing variations in the sampling of the template waveform. Effectively, combination of the basis waveforms allows interpolation for a more perfect fit of the undesired interfering signal. It also allows adaptation to waveforms that vary somewhat in basic shape from the common pattern.

Figs. 2 and 3 depict an example of this type of cancellation. The figures show a 22-minute segment of data from the SPITS array, part of a much longer stream interval plagued by thousands of icequakes. Prior to the processing depicted in the figures our detection framework was run on 7 days of continuous data (Julian days 145-152) to

Approved for public release; distribution is unlimited.

obtain representative waveforms for the icequakes. The system produced 329 detections in 7 groups with waveforms so similar that all were aggregated into a single cluster and used to generate a single rank-3 subspace representation. The representation was used in a subspace detector to detect and time the occurrences of icequakes in the 22-minute segment. Fig. 2 shows the subspace detection statistic on top with waveforms from two channels of the SPITS array below it.

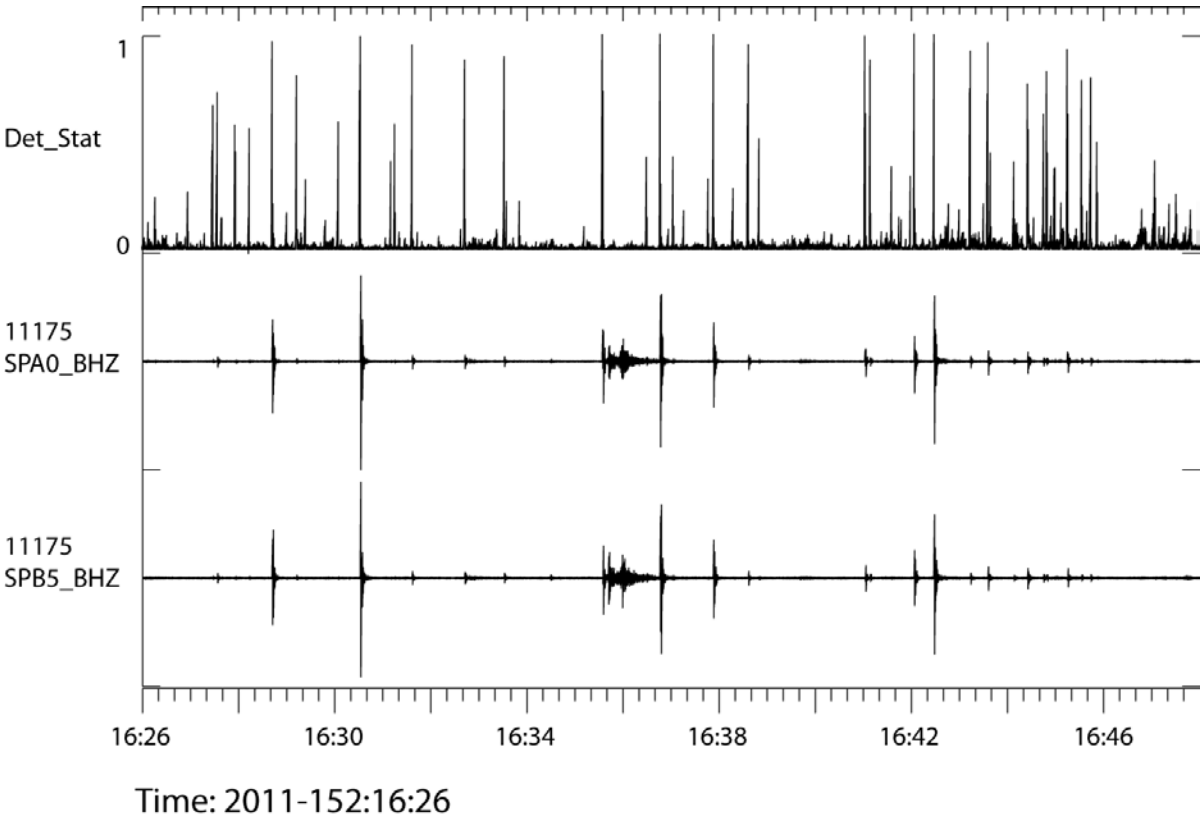

Fig. 2 SPITS vertical component recordings of data from 1 June 2011 *(bottom two traces). The 22 minute time interval contains many transients from very local events, as well as a regular earthquake-type wavetrain arriving around 16:35. The top trace shows the subspace detection statistic used to identify and time interfering local transients.*

Approved for public release; distribution is unlimited.

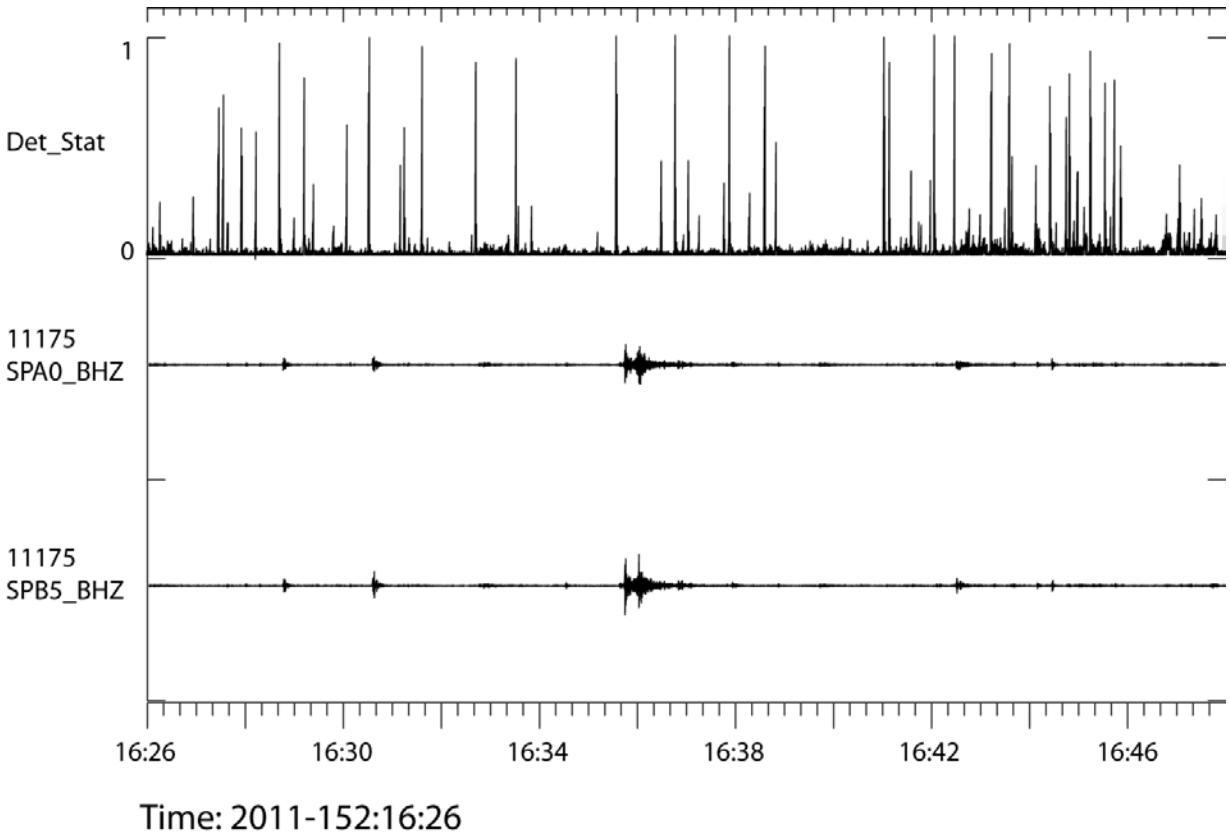

Fig. 3 Data interval of Fig. 2 after cancellation of the local event signals *(bottom two traces). A small fraction of the total energy remains after the cancellation. The residual signals around 16:28, 16:30 and 16:44 correspond to the most energetic of the local events. This figure is plotted to the same amplitude scale as the previous figure.*

Following detection, all events found with a statistic greater than 0.6 were subjected to the cancellation operation outlined earlier, with the result shown in Fig. 3. The top trace of the figure reproduces the detection statistic for reference and the bottom two traces are the residuals after cancellation for the same two stations that were shown in Fig. 2. Except for a small fraction of the energy still remaining after cancellation of the largest signals, the data are efficiently cleaned of the undesired signals, and the wavetrain from a "normal" earthquake arriving around 16:36 stands out clearly. This earthquake is located about 120 km south-east of the array, and has a magnitude of about 1.0. We anticipate that beamforming would substantially suppress the remaining energy in the transients not removed by cancellation.

Practical application of this approach on a routine basis would require a system component like that shown in Fig. 4. This component would maintain an archive of waveform patterns from frequently occurring transients. These patterns would be used in subspace (correlation) detectors operating in a first pass over the data to find occurrences of repeating patterns. The triggers from these detectors would be classified into nuisance transients (local sources of no interest) and triggers from other repeating

Approved for public release; distribution is unlimited.

sources (e.g. mining explosions) that should be noted in a catalog. All triggers above some high threshold would be used to drive cancellation operations in a second pass over the data to produce a “cleaned” data stream that would be passed to later stages of the pipeline for further detection processing. Due to the proximity of local sources to the station, these preprocessing steps could be performed on relatively short windows (e.g. 10 minutes) consistent with the current blocking used in central data acquisition. The triggers also could be used to update waveform templates as desired to allow tracking of evolving local sources.

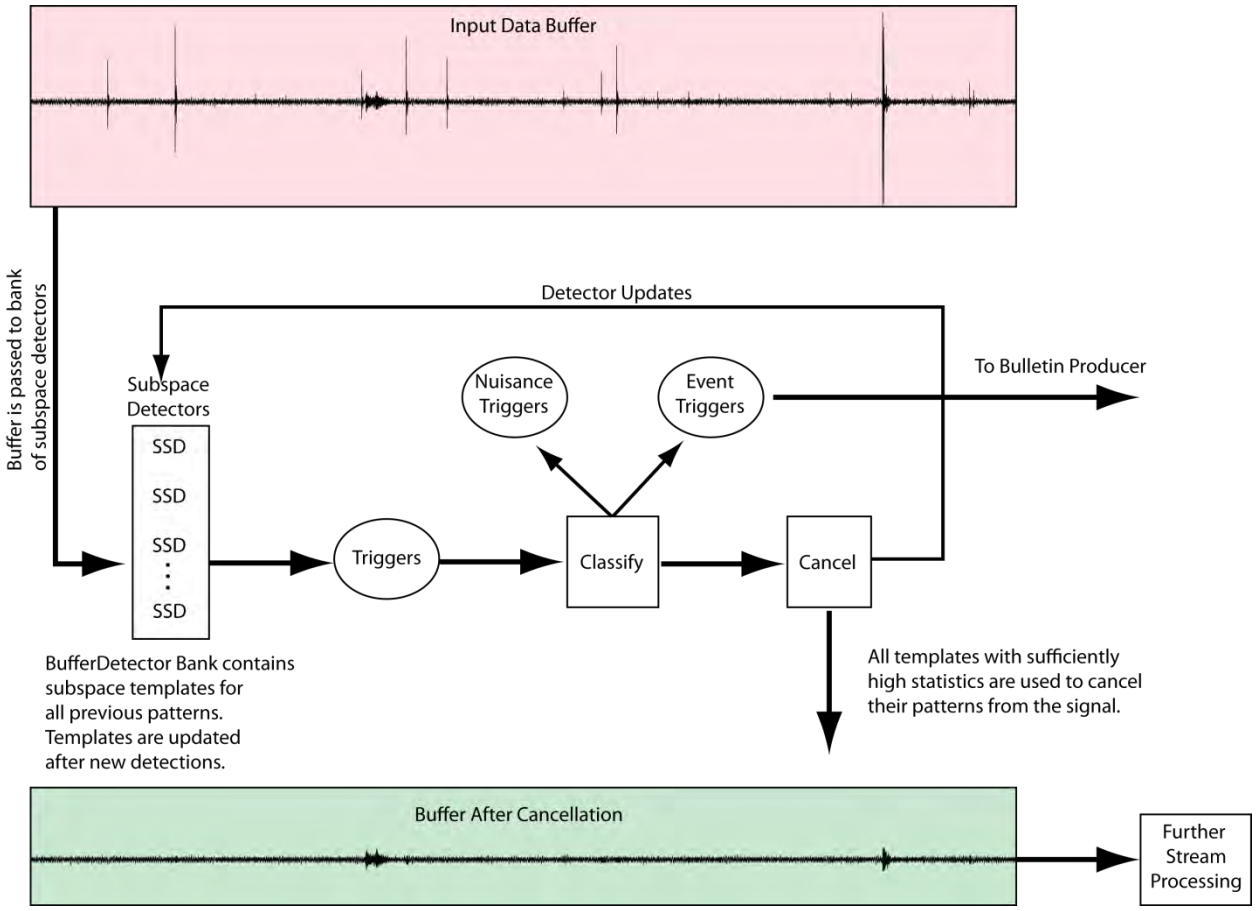

Fig. 4 Block diagram of a system component to remove local transients from a stations data stream. *See text for details.*

Approved for public release; distribution is unlimited.

3.2 Empirical Matched Field Processing and Adaptive Beamforming

The principal thrust of this project is to examine the potential for contextual signal processing to improve detection and association functions in processing pipelines. The idea is to adapt the signal processing algorithms – which today make unrealistically simple assumptions about the seismic background context in which they operate – to make more sophisticated use of present knowledge of context. In support of this idea, one of our objectives is to take a fresh look at adaptive beamforming algorithms for suppression of signals and noise that compete with signals of interest. Adaptive beamforming techniques use estimates of the structure (covariance) of background interference to suppress that interference. These techniques operate by reducing the response of arrays at wavenumbers where interference is strong, while simultaneously maintaining a fixed gain at wavenumbers from which desired (target) signals arrive. These techniques were developed for verification seismology in the 1970s, but were largely abandoned due to their sensitivity to signal model error. They assume that the spatial structures of target signals adhere closely to a plane wave model. Because the spatial structures of incident waves often are significantly non-planar (resulting from unmodeled refraction and scattering), adaptive techniques often suppress target signals as well as interference. In a sense, adaptive methods need to know what signal structure they are protecting as they aggressively configure an array's wavenumber response to suppress interference.

Recent developments in empirical detection methods, especially empirical matched field processing, suggest that signal models may be made sufficiently accurate to allow adaptive beamforming to function as intended. The approach, in keeping with modern trends of large scale waveform archival, is to use past events to provide target waveform calibrations. Empirical matched field processing estimates the spatial structure of target phases from prior observations in a large collection of narrow frequency bands. The quasi-frequency domain calibration that results minimizes sensitivity to variable or unknown source time histories.

As used here, matched field processing is closely related to frequency-domain beamforming, and, in fact, can be viewed as a calibrated form of beamforming. To describe the technique, we first examine a narrowband approximation to frequency-domain beamforming, then generalize the result. We start with the discrete-time samples $\mathrm{r}[\mathrm{n}]$ of the wavefield observed by an array:

Approved for public release; distribution is unlimited.

$$\mathrm{r}[\mathrm{n}] \quad = \quad \begin{bmatrix} \mathrm{r}(\mathrm{x}_1, \mathrm{n}\Delta\mathrm{t}) \\ \mathrm{r}(\mathrm{x}_2, \mathrm{n}\Delta\mathrm{t}) \\ \vdots \\ \mathrm{r}(\mathrm{x}_\mathrm{M}, \mathrm{n}\Delta\mathrm{t}) \end{bmatrix} \tag{1}$$

We treat the observed array signal as a vector process: quantities indicated in bold lower case are vectors and in bold upper case, matrices. Scalar quantities are indicated with an italic font. The vectors $\{\mathrm{x}_\mathrm{i}\}$ represent the locations of the sensors. M is the number of sensors in the array, and each of the i sensors observes the ground motion sequence $\mathrm{r}(\mathrm{x}_\mathrm{i}, \mathrm{n}\Delta\mathrm{t})$. The sampling interval is denoted by $\Delta\mathrm{t}$.

We take a quasi-frequency domain approach in which the data are partitioned into a large number (N) of narrow frequency bands through a bank of bandpass filters:

$$\mathrm{r}_\mathrm{l}[\mathrm{n}] \quad = \quad \sum_\mathrm{k} \mathrm{h}_\mathrm{l}[\mathrm{k}]\, \mathrm{r}[\mathrm{n}-\mathrm{k}] \tag{2}$$

The frequency-domain response of each of the filters is approximately one in a band centered at frequency $\mathrm{l}\Delta\mathrm{f}$ with bandwidth $\Delta\mathrm{f} = 1/\mathrm{N}\Delta\mathrm{t}$ and is zero at all other frequencies. The impulse responses of the filters are related to a baseband lowpass filter $\mathrm{h}_0[\mathrm{k}]$ by complex exponential modulations:

$$\mathrm{h}_\mathrm{l}[\mathrm{k}] \quad = \quad \mathrm{e}^{\mathrm{i}\Omega_\mathrm{o}\mathrm{l}\mathrm{k}}\mathrm{h}_0[\mathrm{k}]\,; \qquad \mathrm{l} = 0, \cdots, \mathrm{N}-1; \qquad \Omega_\mathrm{o} = 2\,\pi\,\Delta\mathrm{f}\Delta\mathrm{t} \tag{3}$$

Defining as $\mathrm{H}_0(\Omega)$ the frequency response (discrete Fourier transform) of the baseband filter:

$$\mathrm{H}_0(\Omega) \quad = \quad \sum_\mathrm{k} \mathrm{h}_0[\mathrm{k}]\ \mathrm{e}^{\mathrm{i}\Omega\mathrm{k}} \tag{4}$$

then the choice of (3) for the family of filterbank impulse responses produces a collection of filters that are frequency translations of the baseband filter:

$$\mathrm{H}_\mathrm{l}(\Omega) \quad = \quad \mathrm{H}_0(\Omega - \mathrm{l}\Omega_\mathrm{o}) \tag{5}$$

It is possible to choose the baseband impulse response in such a way that the summed response of the filters is 1 across the frequency band of interest:

Approved for public release; distribution is unlimited.

$$\sum_{l} H_l(\Omega) \;=\; 1 \tag{6}$$

which allows exact reconstruction of the original signal from its narrowband components. The filters $H_l(\Omega)$ are one sided in the frequency domain, hence complex. Consequently, the narrowband waveforms $r_l[n]$ are complex analytic signals.

In the time domain the beamforming operation consists of shifting operations on each waveform channel followed by summation of the results. In the frequency domain, the operation corresponding to a shift is multiplication of the Fourier transform of the i^{th} waveform by a frequency-dependent complex exponential phase factor

$$e^{-i\,\omega\,\Delta_i} \tag{7}$$

In the plane wave model of propagation, the delay $\Delta_i = \nu \cdot x_i$, where ν is the slowness vector defining (reciprocal) speed and the direction of propagation. The narrowband approximation to the frequency-domain phase shifting product consists of multiplication of each channels complex analytic representation by a phase factor:

$$e^{-i2\,\pi\,l\Delta f\Delta_i} \tag{8}$$

which is constant in band l. The validity of this approximation depends on the bandwidth Δf of the filters being sufficiently small. Sufficiently small is defined by

$$\frac{\Delta f\Delta_i}{2} \ll 1; \;\; \forall\, i \tag{9}$$

so that phase changes over the width of each filter band are insignificant. With this condition, it can be shown that the beam is well-approximated by:

$$b[n] \;=\; \sum_{l} w_l^H\, r_l[n] \tag{10}$$

where the complex weight vector w_l assembles the phase delay factors for all channels:

$$w_l \;=\; M^{-1/2} \begin{bmatrix} e^{-i2\,\pi\,l\Delta f\,\nu\,\cdot x_1} \\ e^{-i2\,\pi\,l\Delta f\,\nu\,\cdot x_2} \\ \vdots \\ e^{-i2\,\pi\,l\Delta f\,\nu\,\cdot x_M} \end{bmatrix} \tag{11}$$

Approved for public release; distribution is unlimited.

The scale factor $M^{-1/2}$ is chosen to assure that $w_l^H w_l = 1$. The vector w_l is known as the steering vector for band l, and the inner product $w_l^H\, r_l[n]$ in (10) implements the narrowband equivalent of the shift and sum operation in that band.

Matched field processing generalizes beamforming by replacing the plane-wave steering vectors of beamforming with steering vectors estimated from prior observations of events at a target location. The estimation procedure consists of performing the narrowband decomposition of the prior observations, then computing covariance matrices for the vector waveforms in each band:

$$R_l = \sum_n r_l[n] r_l^H[n] \tag{12}$$

If more than one event is available as a prior observation, these covariance matrices can be averaged over events. The steering vector for band l is obtained as the top eigenvector in the eigendecomposition:

$$R_l = E_l \Lambda_l E_l^H \tag{13}$$

Assuming ε_l is the eigenvector corresponding to the largest eigenvalue $(\lambda_l)_{max}$, then we choose

$$w_l = \varepsilon_l \tag{14}$$

In the event that the design data consist of a noiseless observation of a pure plane wave, this procedure will reproduce the beamforming steering vector of equation (11). However, if the observed signal departs significantly from a plane wave arriving along the great circle path from source to array, the principal eigenvector or eigenvectors will capture the scattering and refraction details not modeled by a plane wave.

We are examining two options for constructing a statistic $\gamma[n]$ for purposes of declaring detections. The first is an incoherent stack over the narrow bands:

$$\gamma_i[n] = \sum_l \left| w_l^H\, r_l[n] \right|^2 \tag{15}$$

This statistic is the classical wideband detection statistic of matched field processing from the underwater sound community. It estimates the power separately in each band

Approved for public release; distribution is unlimited.

and stacks these estimates incoherently over frequency. Its advantage is that it does not require the steering vectors to correctly preserve phase in the beam across frequencies. Its disadvantage is that it has very poor temporal resolution, roughly the resolution of the impulse response of a single narrowband filter, approximately $1/\Delta f$ seconds.

The second option is a coherent detection statistic, consisting of the squared magnitude of the wideband beam (equation 10):

$$\gamma_c[n] = |b[n]|^2 = \left| \sum_l w_l^H r_l[n] \right|^2 \quad (16)$$

The advantage of this detection statistic is that it potentially has the temporal resolution of a conventional beam. Its disadvantage is that it does require the steering vectors to preserve phase in the beam across the array. In principle, this requirement may be met by referencing the beam to one particular sensor by normalizing the steering vectors to have purely real components for that sensor. In practice, normalization may be hard to achieve if the signal is low or absent in a particular band for that sensor.

We digress at this point to give an example of steering vector calibration. The setting for the example is shown in Fig. 5. We examine the use of one Myanmar event (2007 5/16 8:56:16.0, M_w 6.3; 20.47°N 100.69°E) as a calibration for a second event occurring nearly 4 years later (2011 3/24 13:55:12, M_w 6.9; 20.69°N 99.82°E). Both events are large and were recorded by the ILAR array with high SNR at many frequencies (see waveforms in Fig. 6). We followed the design procedure just outlined for estimating steering vectors from waveforms of the 2007 event, using $N = 512$ bands. Since the data are real, only 257 of these bands are independent, and it suffices to estimate steering vectors in just the bands with indices $l = 0, \ldots, 256$. The bandwidth for the narrowband filters in this analysis is 20/512 = 0.003904 Hz (the data are sampled at 20 sps). The covariance matrices were computed over windows containing the first 25 seconds of the P arrival, for both the design event and the later target event.

Approved for public release; distribution is unlimited.

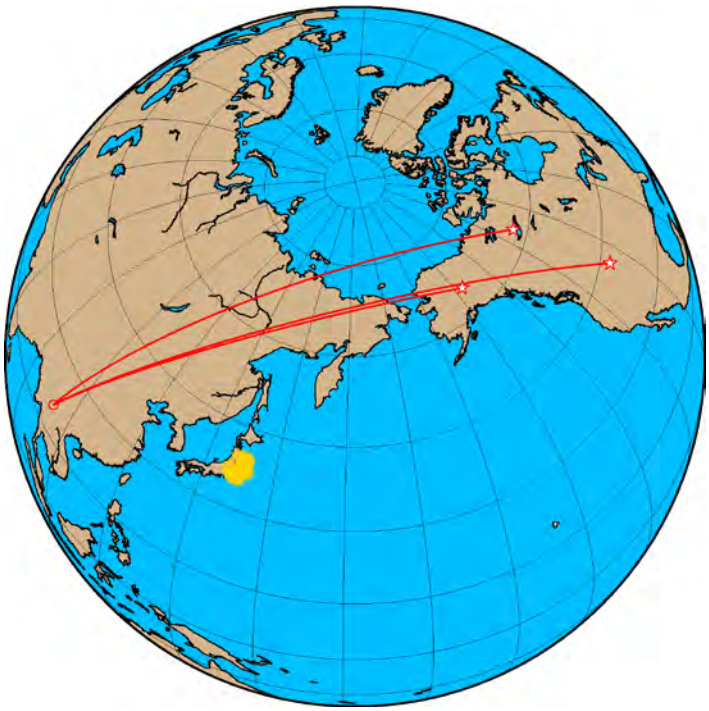

Fig. 5 Teleseismic paths from earthquakes in Myanmar to three North American arrays. *The path length to ILAR (the nearest array) is about 8950 kilometers. Tohoku aftershock cloud shown in yellow.*

Approved for public release; distribution is unlimited.

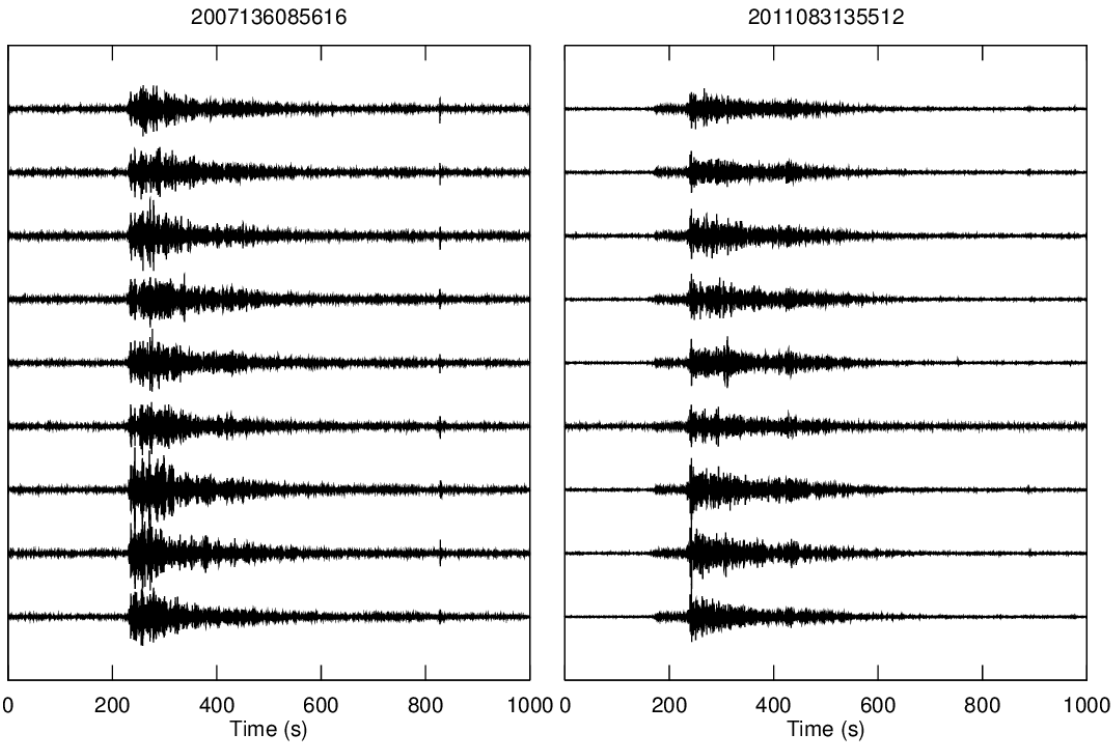

Fig. 6 Waveforms of Myanmar calibration event (left) and target event (right), recorded at ILAR. Waveforms from nine stations (IL01-IL09) are shown. *Note that the second event is closely preceded by a smaller event, perhaps a Tohoku aftershock. The signals have been filtered into the 1-3 Hz band.*

Approved for public release; distribution is unlimited.

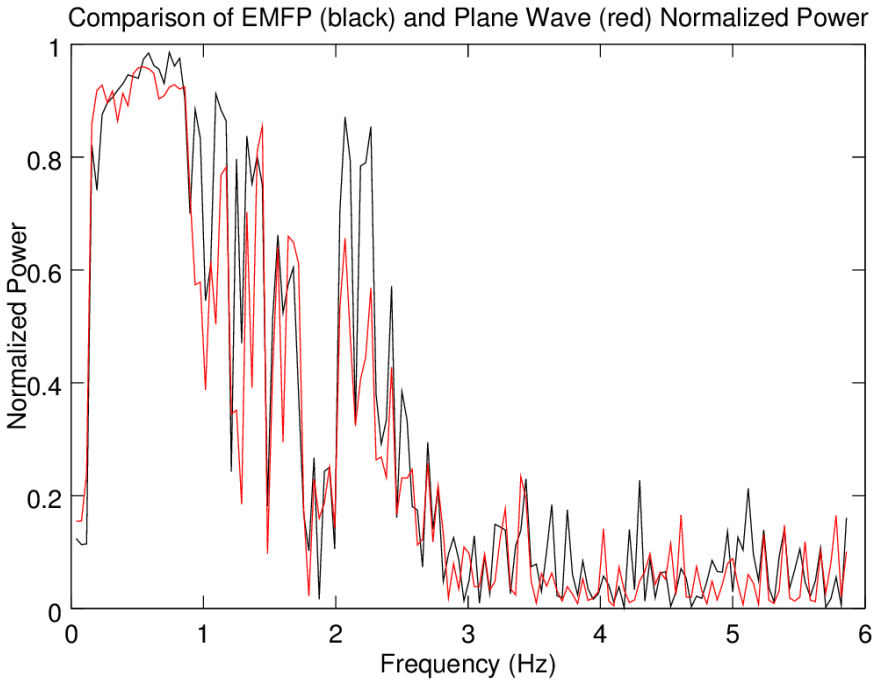

Fig. 7 Comparison of empirical matched field processing (black) and plane wave (red) power for the 2011 3/24 M_w 6.9 event.

A comparison of beamforming performance for plane wave and empirical matched field steering vectors is shown in Fig. 7. The measurements shown in this figure are of normalized power:

$$\frac{w_l^H R_l w_l}{tr\{R_l\}} \tag{17}$$

The plane-wave steering vectors (equation 11) were determined by the great-circle backazimuth (299.2 degrees) from the array to the source and by the P velocity predicted by IASP-91 for the epicentral distance 8950 km. We see that below 2 Hz, the plane wave and the empirical steering vectors are performing nearly equally well – sometimes one is better, sometimes the other; values are comparable. Except in the vicinity of 1 Hz, where the empirical steering vectors appear to have a distinct advantage. Above 2 Hz, the empirical steering vectors perform substantially better than the theoretical plane wave vectors. This observation is consistent with expectations that scattering and refraction – driven by heterogeneities in the propagation medium – should increase with increasing frequency.

Approved for public release; distribution is unlimited.

Beamforming and matched field processing are optimum estimators of the signal (and optimum detectors) in the case that the background noise is white, uncorrelated among sensors and of equal power on all sensors. In the event that the noise is correlated among sensors, the optimum choice for steering vectors is the adaptive beamformer weighting [Capon et al., 1967] also known as the minimum variance, distortionless receiver (MVDR):

$$w_l = \frac{R_{nl}^{-1} \varepsilon_l}{\varepsilon^H R_{nl}^{-1} \varepsilon_l} \quad (18)$$

To obtain these weights, an estimate of the noise covariance R_{nl} is required, which means that a separate sample of noise data must be available for estimating covariance. In the example we show later, the "noise" will, in reality be non-stationary aftershock observations. The theory of adaptive beamforming was developed for stationary processes, but our example shows that it works to a degree for strongly non-stationary processes such as interfering aftershocks. In any case, the MVDR weighting has the effect of altering the wavenumber response of the array to suppress wavenumbers corresponding to strong noise components, while passing a desired signal with spatial structure defined by the steering vector ε_l.

3.3 Evaluation of Seismic Event Hypotheses

In processing pipelines for detecting and locating seismic events, phase detections at different stations are associated if there an event hypothesis capable of explaining the observation of each of the phases at the appropriate times. An event hypothesis may be based on a small number of arrivals which, while technically consistent with an event hypothesis at a given location and a given time, would not be expected to be the only observations seen. If an event hypothesis is missing detections from key stations which would be expected to detect signals, and yet includes phases at stations for which the likelihood of detection is very low, the hypothesis is likely to be incorrect. To address this issue, NORSAR has over some time assessed detection capabilities for IMS stations for application to the global phase association process (Kværna and Ringdal, 2013). The basis for this work is the events and phase readings reported in the IDC reviewed Event Bulletin (REB), as well as the corresponding lists of non-detecting stations. One type of results from this study is a map of regionalized capability estimates for each IMS station (see Fig. 8). To provide station detection capability estimates for regions with no, or very few, calibration events, a standard amplitude-distance curve (e.g. Murphy and Barker, 2003) is fitted to regionalized estimates.

Approved for public release; distribution is unlimited.

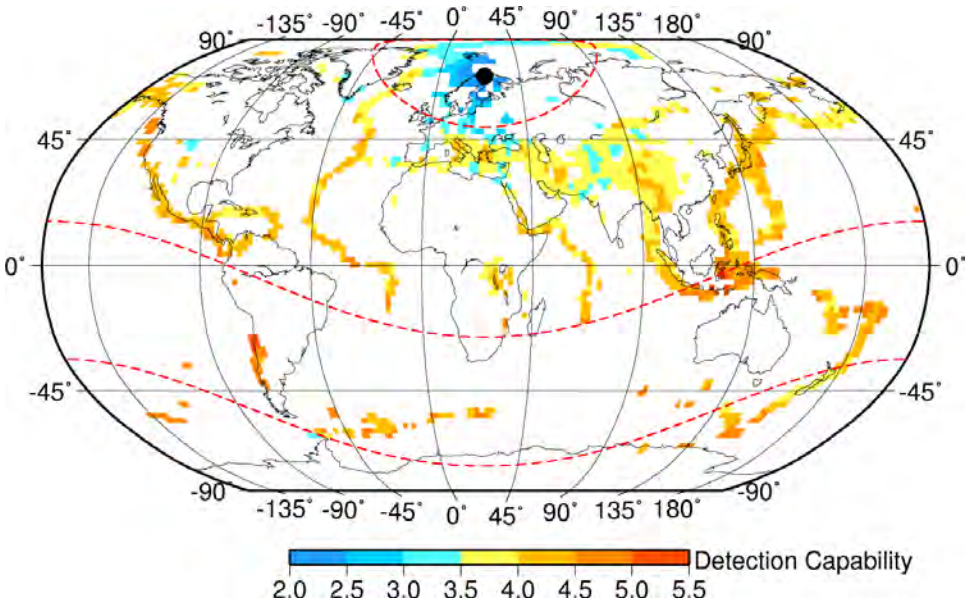

Fig. 8 P-phase detection capability estimates for the ARCES array in 2°x2° bins. *The red stippled curves denote distances of 20, 95 and 144 degrees from ARCES. The blue bins denote "bright spots" where the ARCES detection capability is particularly good (* $m_b \leq 3.5$*).*

Approved for public release; distribution is unlimited.

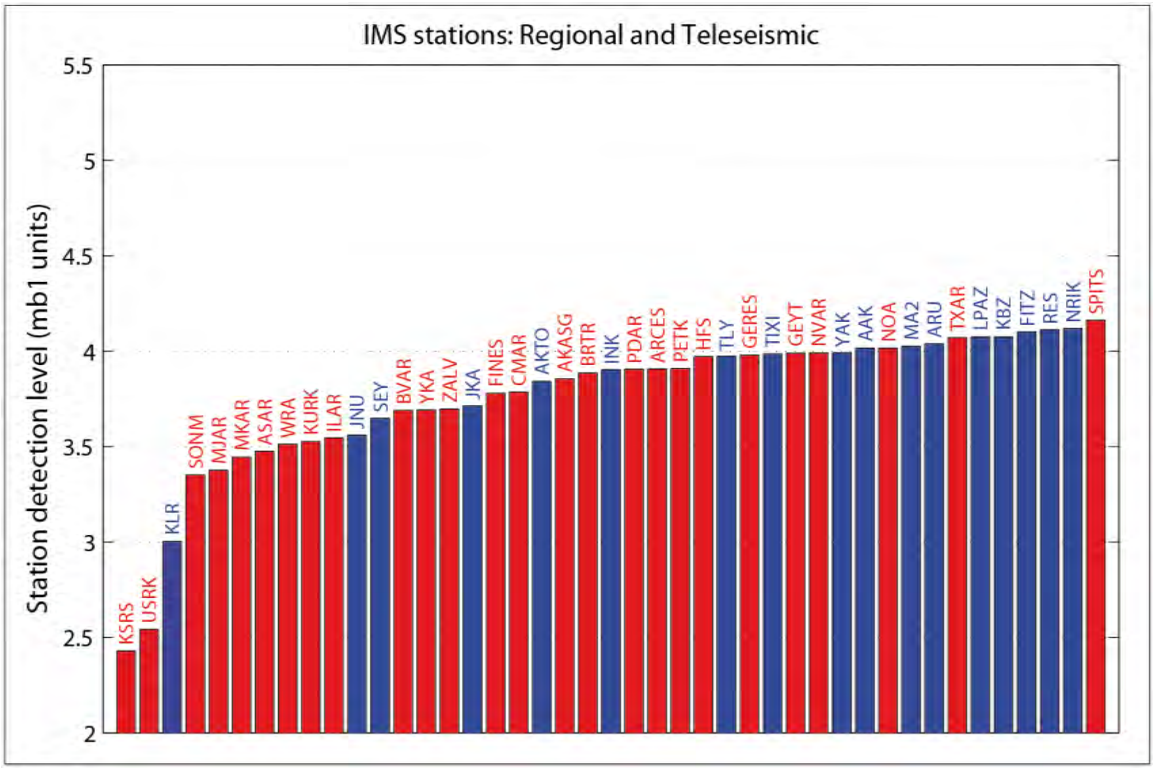

Fig. 9 Predicted detection thresholds in units of magnitudes for the 44 IMS stations (both array and 3-component, primary and auxiliary) with best predicted detection capability for the North Korea nuclear test site (assumed location 41.28°N, 129.07°E). *Array stations (both primary and auxiliary) are displayed in red and 3-component stations (both primary and auxiliary) are displayed in blue.*

These types of results can provide quantitative information about which stations are most likely, or most unlikely, to detect signals from events in any given source location. An example is shown in Fig. 9 for a source region surrounding the North Korea nuclear test site (41.28°N, 129.07°E). Taking the North Korea test site as an example, we can, based on several years of REB bulletin data, estimate the likely detection capability at a given set of stations using the approach of Kværna and Ringdal (2013). Fig. 9 displays the 44 best-case detection capability estimates for this test site for IMS stations (with arrays and 3-component stations differentiated) and Fig. 10 displays the detection capability estimates for this set of stations as a function of epicentral distance. We find that these station detection capability estimates are in good agreement with the distribution of signal-to-noise ratios reported in the REB for the North Korean underground tests. The locations of the stations are displayed relative to the source region in Fig. 11.

Approved for public release; distribution is unlimited.

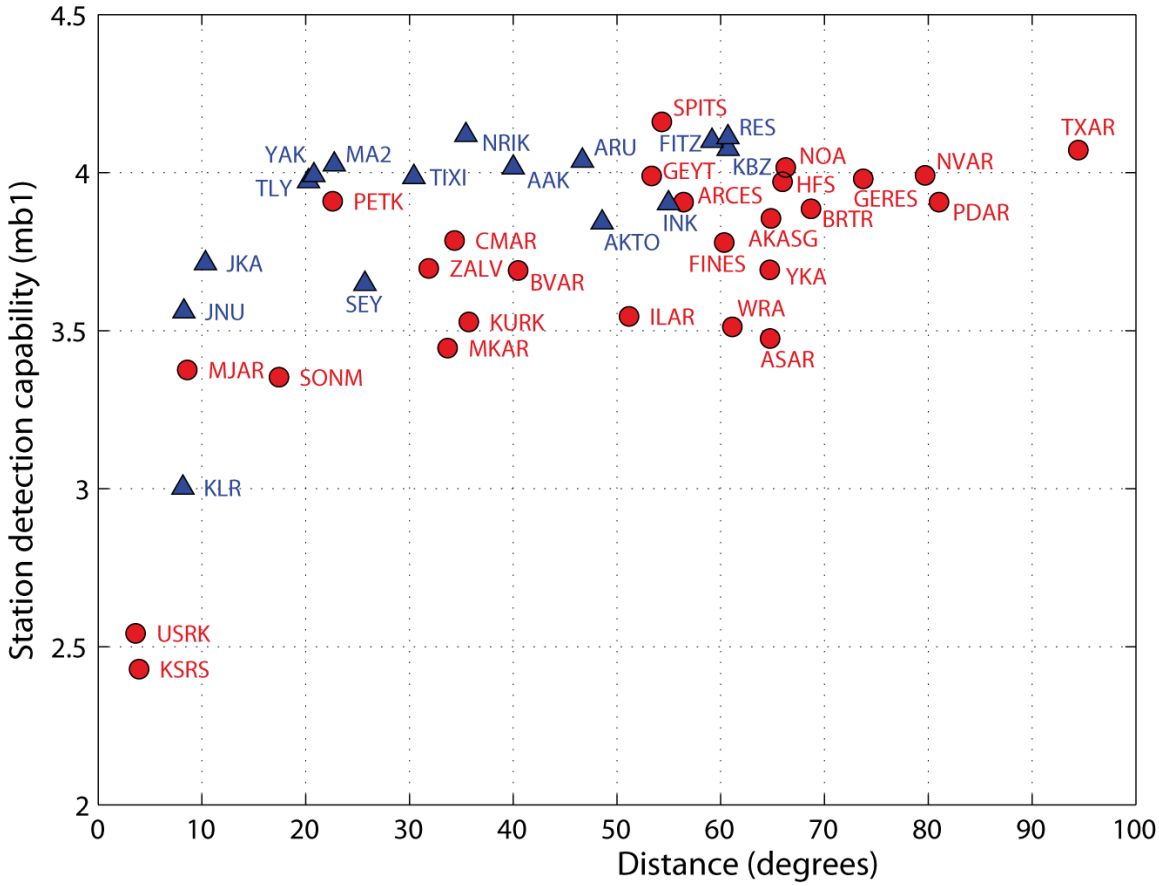

Fig. 10 Predicted detection thresholds in units of magnitudes for the North Korea nuclear test site (assumed location 41.28°N, 129.07°E) as a function of epicentral distance for the 44 IMS stations displayed in Fig. 9. *The array stations are displayed using red circles and the 3-component stations are displayed using blue triangles.*

Approved for public release; distribution is unlimited.

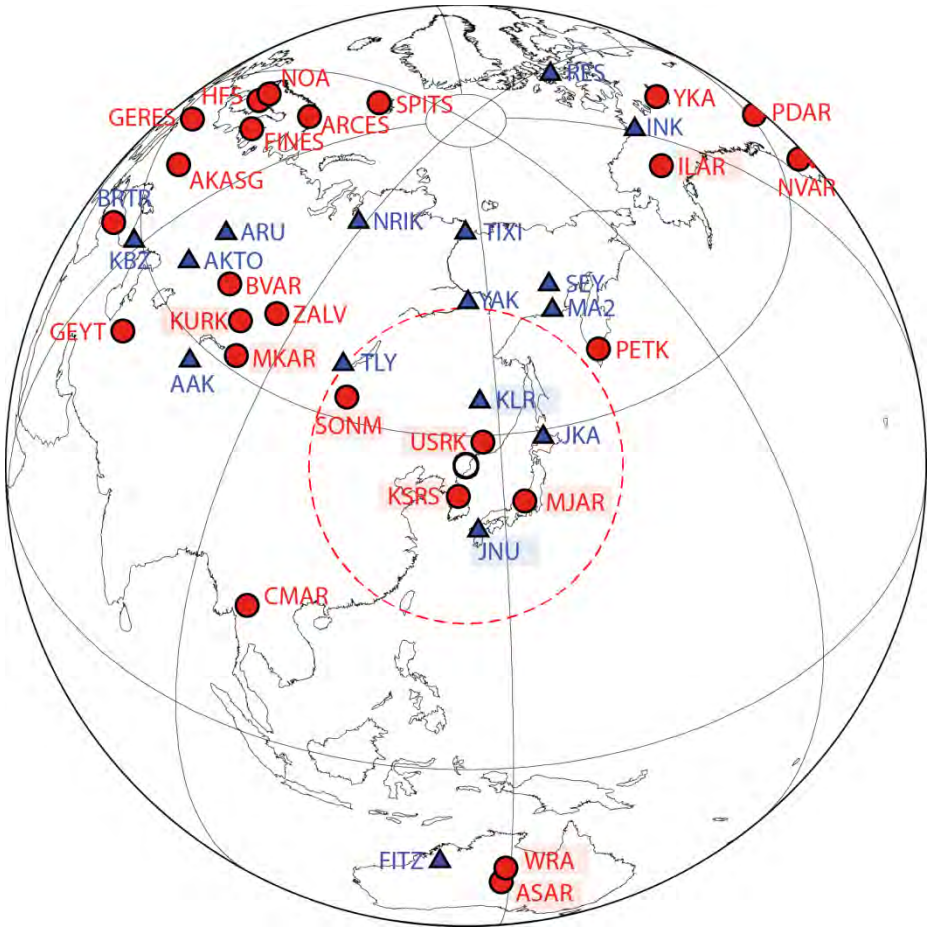

Fig. 11 Locations of the stations displayed in the previous two figures. *The most sensitive stations for this source region have the station codes shaded.*

The primary goal of this study is to assess the consistency of the automatic phase associations of the IDC SEL's, and to improve the performance of the phase association algorithm. However, the information about the station detection capabilities, as shown in Fig. 9, can equally well be used as a tool for recall data processing of stations having a high likelihood of observing signals from a hypothesized event. Similarly, recall data processing can be done to reject incorrectly associated phases at stations having a low probability of detecting the event.

It is the intention to exploit this information in several ways:

1) to predict the stations that should be severely affected by a developing aftershock sequence,
2) to identify the stations that should observe certain phases for an hypothesized event, but failed to do so, and

Approved for public release; distribution is unlimited.

3) to identify situations where stations reported defining phase detections (those used to build an event), but were unlikely to observe phases for the hypothesized event.

These are all situations where supervisory function in a future pipeline might initiate tailored processing or reprocessing of waveform data. In the first case, beamforming algorithms might be replaced for the duration of the aftershock sequence with adaptive methods suitable for suppressing aftershock waveforms. In the second case, specialized beams might be deployed to search for the missing phases. And in the third case, specialized beams might be used to confirm or refute the existence of defining phases at the reporting stations. This last process would serve as a check on improperly associated phase detections. In all three cases, adaptive beamforming algorithms might be used to improve on results obtained by conventional (recipe) beamforming algorithms in a first pass over the data.

A context-driven processing framework will attempt to evaluate how consistent candidate event hypotheses are with the observations. Subsequent reprocessing of the data stream(s) will depend upon the event hypotheses which are considered by the previous iteration to best represent the true event history. For this we need an empirical (or semi-empirical) means of predicting the likely observations from a given hypothetical event history. A module which predicts the likelihood of a set of observations from a given event hypothesis is being developed. This module can generally be used as a tool to find the optimal set of stations for monitoring and detecting events at any geographical location.

Approved for public release; distribution is unlimited.

4 RESULTS AND DISCUSSION

Each of the following three subsections considers in turn the progress made during the first 12 months of this contract on the material described in subsections (3.1), (3.2), and (3.3).

4.1 Subspace Signal Cancellation

The framework developed has been extended to support a variant of subspace detector whose detections are used exclusively for signal cancellation. Currently, there is no provision to create these detectors automatically as part of the framework execution, although that is an eventual goal. Instead, these detectors are created in the Builder program in the same manner as other detectors, but with the detector type set to "CANCELLATION_SUBSPACE". Fig. 12 displays a screenshot from the Builder program and shows a set of signals being used to create a cancellation detector. The shaded area indicates the time span that will be used by the new detector.

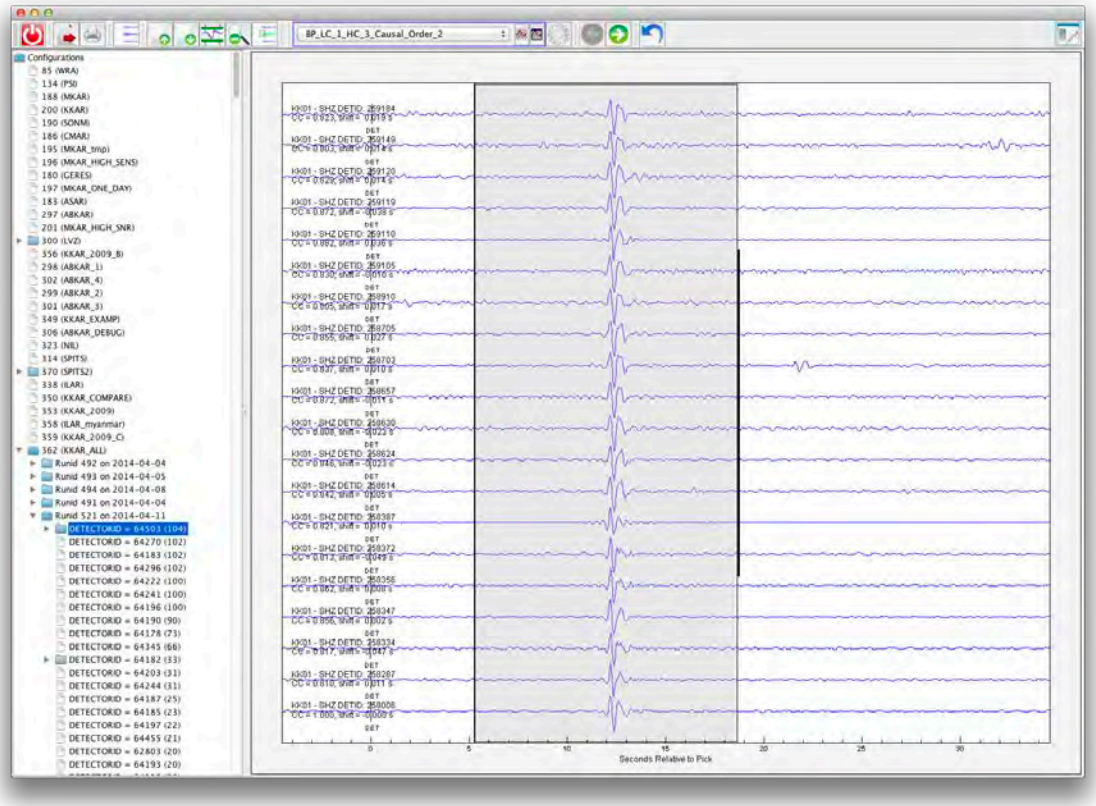

Fig. 12 The preparation of a cancellation detector in the Builder program.

Approved for public release; distribution is unlimited.

Within the framework, the cancellation detectors are kept in a different collection from the other detectors. After each new block of data is pre-processed (band pass filtered and decimated) the cancellation detectors are applied in the same manner as ordinary detectors. The normal trigger reconciliation process is used to find the “best” of any coincident triggers, and the triggers are written into the database.

The cancellation data are produced by doing a least squares fit of the subspace basis vectors of the best detector to the portion of the stream that produced the maximum value of the detection statistic. These data are simply a channel-by-channel best-approximation of the signal which are then subtracted from the pre-processed data. An example of cancellation is shown below in Fig. 13. The black traces show a single channel prior to cancellation, and the red traces are the same channel after cancellation. The cancellation was rather effective on the first and 4th signals, but performed less well on the other two signals. This may be due to the rank of the detector. (In this case the cancellation detector, although formed from 6 detections, was only rank 1.)

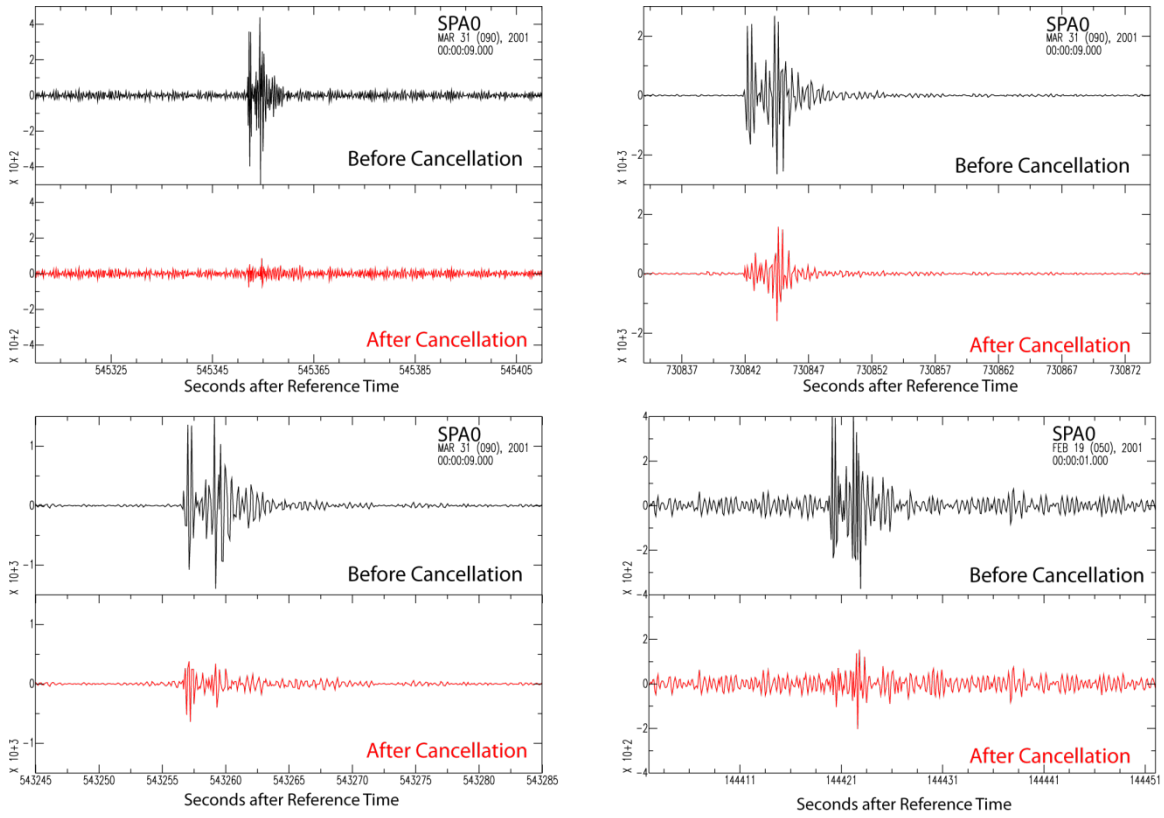

Fig. 13 The result of signal cancellation on four signals on the vertical component of the center element of the SPITS array.

Approved for public release; distribution is unlimited.

Clearly some experimentation is necessary to determine a set of “best practices” in creation of cancellation detectors. However, a more important task at this point is making modifications to the framework architecture to support cancellation (and other forms of advanced pre-processing) in a more “pluggable” manner. The current framework architecture has rather extensive coupling (through shared data structures) of the detectors, the stream processor, and the pre-processor. This was done in the interest of performance and memory conservation. But a side effect of the coupling is that it makes it much more challenging to insert new functionality into the processing flow. Also, we are beginning to plan for deployment of the framework on a large cluster. In that context, there will no longer be shared memory, so communication between objects must be by messaging. Because of this, the next stage of development will be to refactor the framework code into a new message-based architecture. Fig. 14 shows the processing flow through the new architecture and the way in which cancellation will fit into the processing.

Approved for public release; distribution is unlimited.

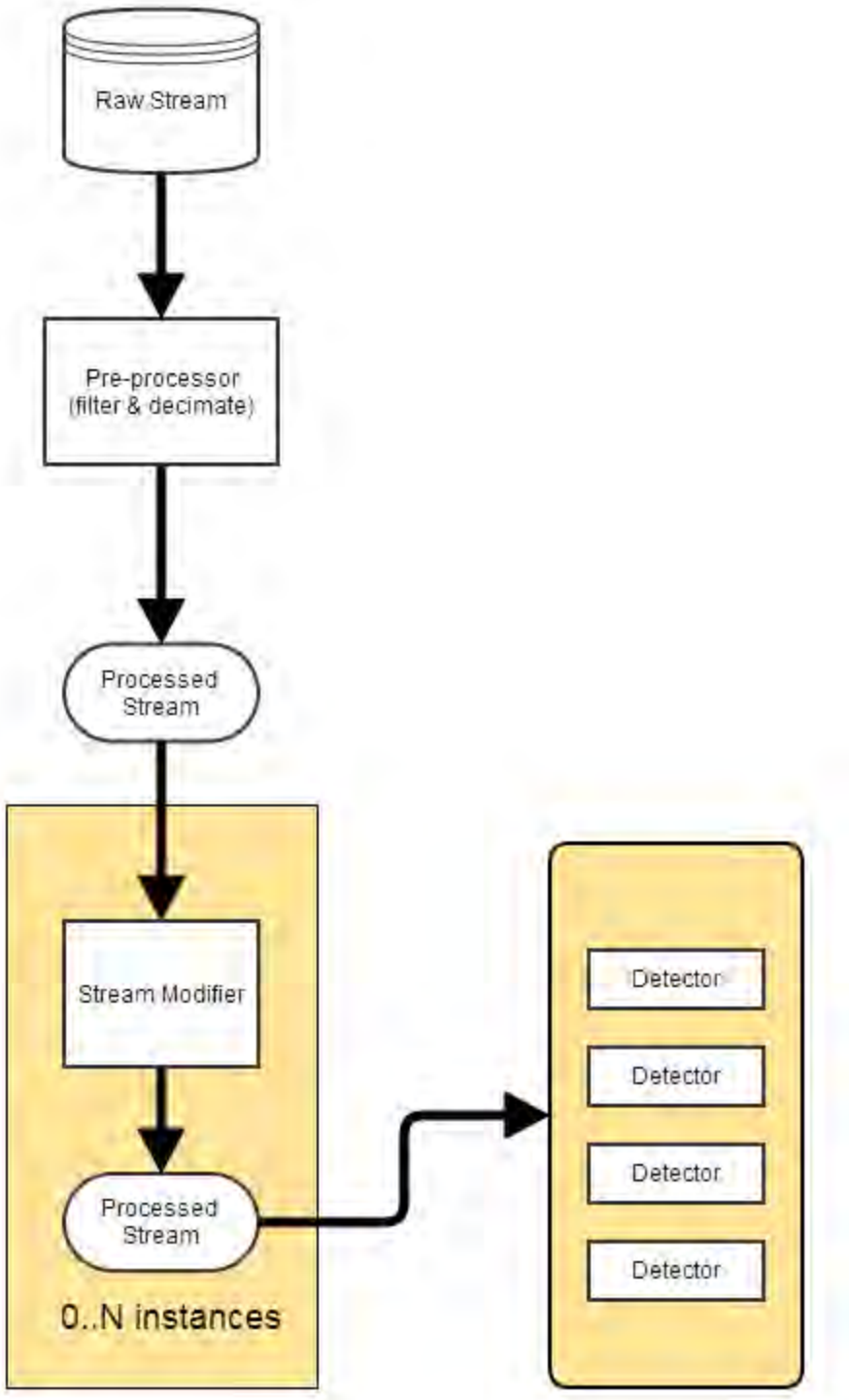

Fig. 14 The processing flow through the new message-based architecture.

Approved for public release; distribution is unlimited.

In Fig. 14, the rectangular boxes represent classes that act on data, and the ovals represent message classes. The StreamModifier is an optional component that takes as input a ProcessedStream and that emits a ProcessedStream. Both StreamModifier and Detector both extend an interface called ProcessedStreamConsumer. This is shown in Fig. 15.

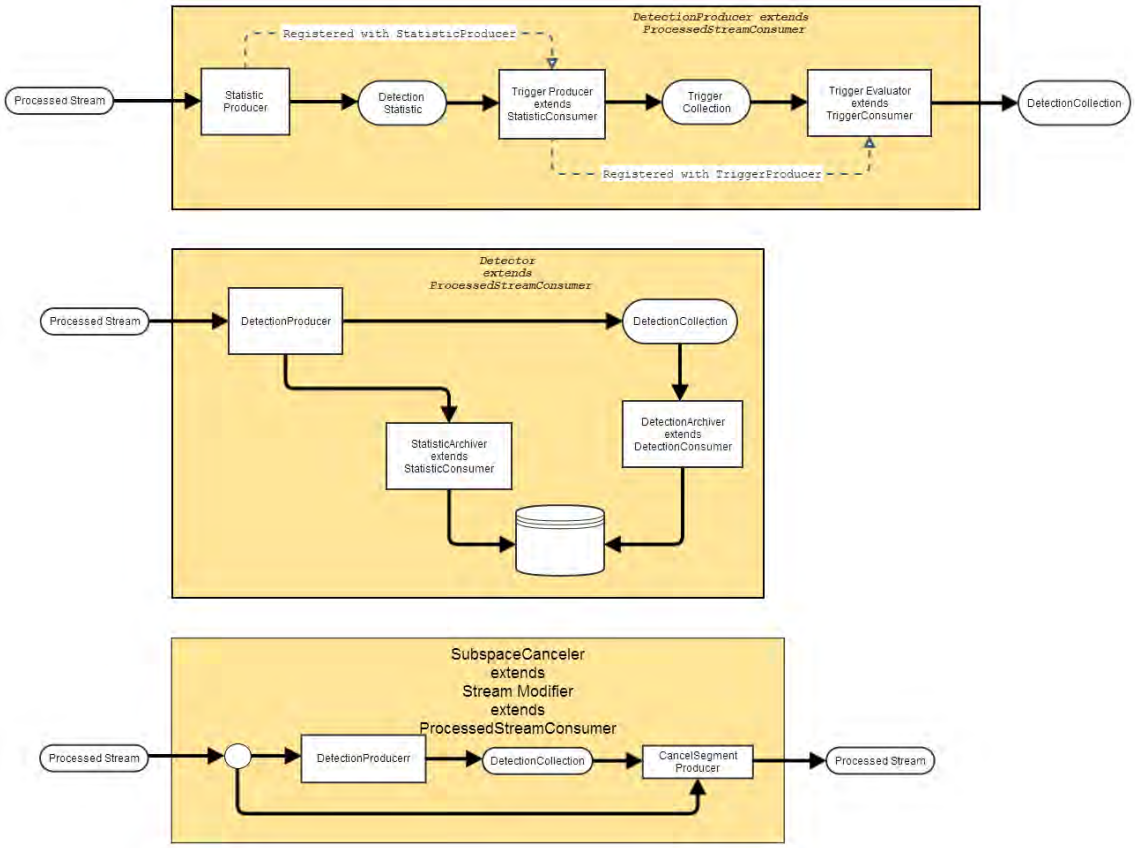

Fig. 15 Details of the components referenced in Fig. 14.

All of these components displayed in Fig. 15 implement the interface ProcessedStreamConsumer and can be registered as consumers for any object that emits ProcessedStream objects. The DetectionProducer class produces a DetectionCollection. Although it does not archive any objects, it will accept listener classes that can archive detection statistics and Triggers.

The Detector class is mainly a wrapper around DetectionProducer that can archive its products. The SubspaceCanceler also wraps a DetectionProducer, but it uses the DetectionCollection to drive the creation and application of CancellationSegments. These are subtracted from the input ProcessedStream which is the output of this object.

Approved for public release; distribution is unlimited.

4.2 Empirical Matched Field Processing and Adaptive Beamforming

In this section, we give an example of plane wave and matched field processing for the incoherent case (equation 15), both for the "conventional" (14) and adaptive (18) processors. The objective is to demonstrate improvements in beamforming performance in a spotlight situation during a major aftershock sequence. The example is contrived by burying the 2006 North Korean nuclear test among aftershocks of the 2008 Sichuan (Wenchuan) earthquake at a 1:1 ratio using data from the 19-channel ASAR (Alice Springs) array. Fig. 16 shows the geographic configuration of the example. Fig. 17 shows two channels of the superimposed ASAR data (sampling frequency: 20 sps). At its natural amplitude, the signal from the test is much smaller than many of the aftershock signals.

Fig. 16 Geographic configuration of the test with superimposed Wenchuan aftershocks and the 2006 DPRK nuclear test. *The North Korean test site is shown with the black outlined star and the observing array (ASAR) is shown with the red outlined star. The Wenchuan sequence aftershocks for the first day (May 12, 2008) are shown in white.*

The results of processing the array data with four different matched field processing algorithms is shown in Fig. 18. The figure shows detection statistics for 19,000 seconds of data, starting 6,000 seconds from a reference point (the main Sichuan event); signals from the North Korean test were buried approximately 18,000 seconds into the record from the main shock.

Approved for public release; distribution is unlimited.

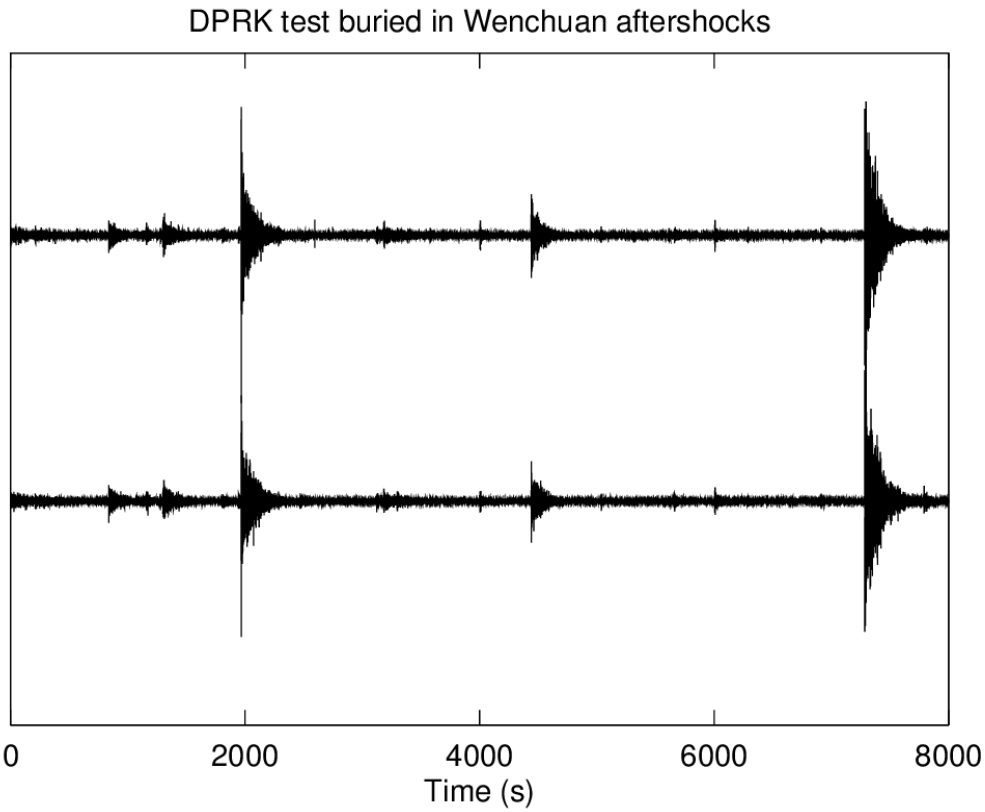

Fig 17 2006 DPRK test superimposed at natural amplitude among aftershocks of the 2008 Wenchuan sequence. *Waveforms from two channels (AS01, AS10) of the 19-channel ASAR array are depicted. The signal from the test appears at 6000 seconds in this plot. The data have been filtered into the 1-3 Hz band.*

The top trace in Fig. 18 displays the power output (equation 15) of an empirical matched field processor operating in the 1-3 Hz band (all four examples are in this band and all traces are plotted to the same scale). Waveforms from the 2009 North Korean tests observed at ASAR were used to obtain the signal model. The second trace shows the power output of a matched field processor which used the best plane-wave model for the wavefield incident from the North Korean test site. Comparing these two traces it is apparent that the empirical processor captures about twice the power from the desired signal as that captured by the plane wave processor. The empirical processor also suppresses the aftershocks slightly better than the plane wave processor.

The bottom two traces show the corresponding power traces for the empirical (third trace) and plane-wave (fourth trace) MVDR matched field processors. MVDR processors require an estimate of the background noise covariance, which in this case was supplied by aftershocks in the first 6,000 seconds of data (prior to the displayed traces). Both processors have done an impressive job suppressing aftershock power, but the emprical

Approved for public release; distribution is unlimited.

processor again has captured about twice the energy from the North Korean test compared to the plane-wave processor. Although it is less apparent, the empirical processor has done substantially more to suppress aftershock power than the plane-wave processor with reference to the non-MVDR case.

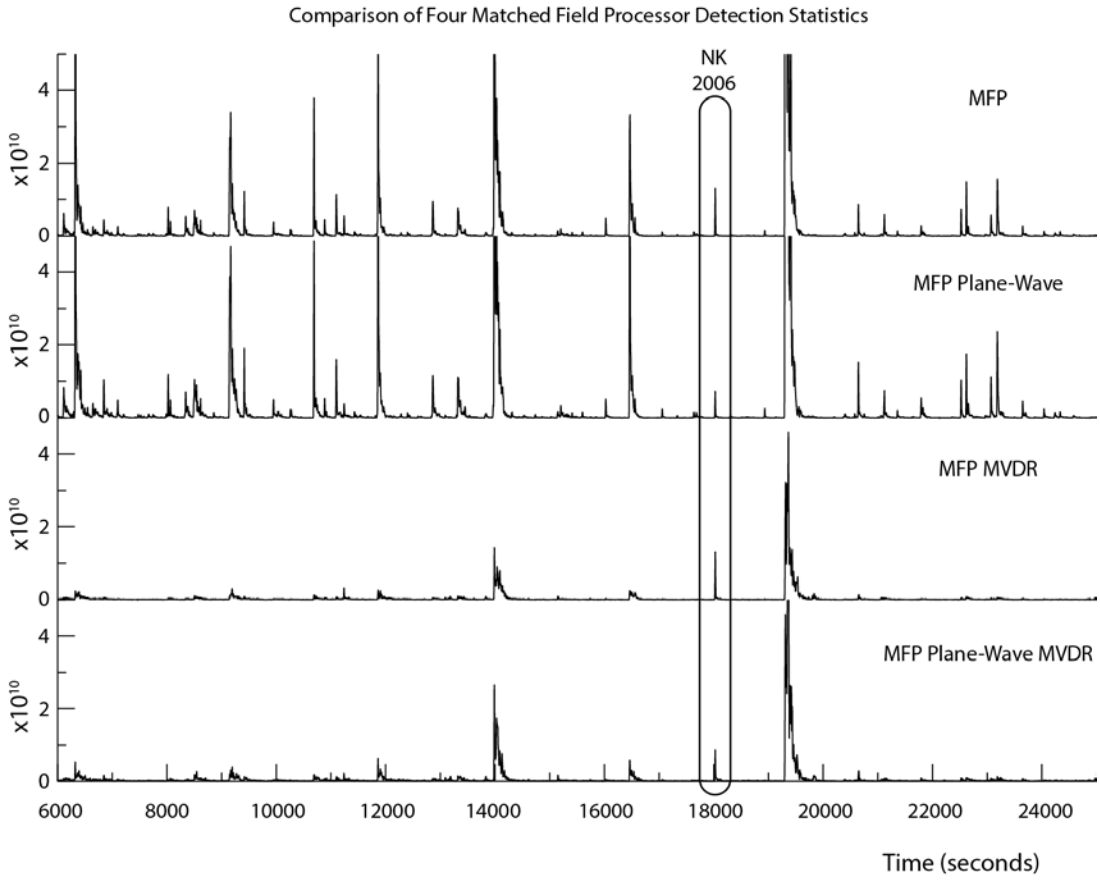

Fig. 18 Detection statistics for four matched field processors operating on data from the ASAR array show that empirical matched field processing with MVDR sidelobe suppression improves detectability of a weak signal among strong aftershocks from a different location.

We anticipate that more dramatic improvements (comparing empirical MVDR to plane-wave MVDR matched field processing) will be obtained with regional observations in higher frequency bands than shown here. The effects of heterogeneities on the signal model are more severe at regional distances.

Approved for public release; distribution is unlimited.

4.3 Testing of Event Hypotheses

A significant component of developing the context-based detection framework is to evaluate approaches for accepting or rejecting phases attached with event hypotheses. Comparing the (fully automatic) SL3 event bulletin and the (reviewed) REB from the International Data Center (IDC) of the Comprehensive Nuclear-Test-Ban Treaty Organization (CTBTO) indicates a number of circumstances in which false event hypotheses have been formed from genuine phase detections and true events which have had to be built manually.

Fig. 19 shows an example of such a case. Two events in central Asia have generated waveforms at stations globally that overlap. The result is that the first event, a deep earthquake in the Hindu Kush region of Afghanistan, is well classified automatically with a qualitatively correct solution in the automatic SL3 bulletin. The second event, a shallow earthquake close to the epicenter of the M=7.6 October 8, 2005, Kashmir event, generates signals which are associated with a spurious event hypothesis.

```
EVENT  3441778 HINDU KUSH REGION, AFGHANISTAN
   Date       Time        Err   RMS Latitude Longitude  Smaj  Smin  Az Depth   Err Ndef Nsta Gap  mdist  Mdist Qual   Author      OrigID
2005/11/05 23:24:27.35   6.75  0.92  36.8415   70.7943  62.7  21.9 176 257.8  54.7   13   13 231  13.11  80.94 m i uk IDC_SEL3:  3441904

Magnitude  Err Nsta Author      OrigID
mb     3.5 0.2   10 IDC_SEL3:  3441904

Sta     Dist  EvAz Phase       Time      TRes  Azim AzRes   Slow   SRes Def   SNR       Amp   Per Qual Magnitude    ArrID
MKAR   13.11  37.1 P        23:27:23.175  -1.2 223.4  -1.4   17.1    4.8 TAS  53.9       0.9  0.33 a__            26110551
BVAR   16.19 359.1 P        23:28:00.437   1.0 158.8 -20.1   12.4    1.5 TAS  49.8       2.6  0.33 a__            26112040
AKTO   16.40 329.9 P        23:28:02.375   0.5 137.6  -3.5   10.9   -0.0 TAS  33.7       1.6  0.33 a__            26111681
SONM   28.28  55.9 P        23:29:57.560   0.5 270.3  10.0    9.1    0.2 TAS  10.3       1.0  0.59 a__ mb     3.3 26110501
BRTR   29.17 287.1 P        23:30:04.550  -0.5 100.8  17.1    8.2   -0.6 TAS   8.5       1.1  0.63 a__ mb     3.3 26110565
AKASG  32.60 308.5 P        23:30:33.900  -0.8  80.6 -18.4    7.8   -0.8 TAS   6.7       0.4  0.35 a__ mb     3.1 26110563
MLR    34.47 298.8 P        23:30:53.384   2.4 169.4  81.0    3.3   -5.2 TAS   4.9       2.4  0.88 a__ mb     3.4 26111758
FINES  37.10 325.9 P        23:31:12.700  -0.2 114.6  -3.8   10.6    0.0 TAS  12.0       2.0  0.67 a__ mb     3.5 26110539
ARCES  40.76 337.5 P        23:31:43.250   0.3 111.7  -1.6    7.5   -0.2 TAS  29.2       2.6  0.44 a__ mb     3.8 26110680
HFS    42.70 321.7 P        23:31:57.994  -0.7  97.3   0.7   10.2    2.2 TAS  15.5       1.8  0.37 a__ mb     3.7 26111908
NOA    44.01 322.8 P        23:32:08.350  -0.8  96.5   0.2    7.7   -0.0 TAS   7.6       1.6  0.61 a__ mb     3.4 26110692
INK    73.61   9.2 P        23:35:32.375   0.4 328.7 -11.1    4.9   -0.9 TAS  11.8       1.8  0.61 a__ mb     3.7 26112063
YKA    80.94   2.5 P        23:36:11.945  -0.6 348.9  -3.4    5.4    0.0 TAS   6.0       0.6  0.34 a__ mb     3.5 26110683
▌

EVENT  3441926 LAKE ISSYK-KUL REGION
   Date       Time        Err   RMS Latitude Longitude  Smaj  Smin  Az Depth   Err Ndef Nsta Gap  mdist  Mdist Qual   Author      OrigID
2005/11/05 23:25:32.57   2.67  0.11  42.6983   79.7545  61.0  33.1  68   0.0f        3    3 203   4.48  80.05 m i uk IDC_SEL3:  3441929

Magnitude  Err Nsta Author      OrigID
ML     2.5 0.4    2 IDC_SEL3:  3441929
mb     4.1        1 IDC_SEL3:  3441929

Sta     Dist  EvAz Phase       Time      TRes  Azim AzRes   Slow   SRes Def   SNR       Amp   Per Qual Magnitude    ArrID
MKAR    4.48  22.9 Pn       23:26:43.725  -0.1 189.5 -15.2   14.8    1.0 TAS   4.6      0.02  0.33 a__ ML     1.4 26110550
BVAR   12.08 332.0 Pn       23:28:24.062  -0.1 167.5  22.4   13.7   -0.0 TAS   3.9       0.4  0.33 a__ ML     3.6 26112161
WRA    80.05 128.9 P        23:37:44.259   0.1 320.4 -10.8    5.2    0.2 TAS   6.6       1.3  0.79 a__ mb     4.1 26110688
WRA    80.05 128.9 tx       23:37:57.384       320.4 -10.9    5.0        ___   7.9       0.5  0.68 a__            26110689
WRA    80.05 128.9 tx       23:38:02.559       321.3 -10.0    5.0        ___   7.9       0.6  0.63 a__            26110690
```

Fig. 19 Extract from SL3 automatic event bulletin from the International Data Center (IDC).
Two event hypotheses are listed here. The first of these corresponded well with an event which was subsequently reviewed and accepted. The second event hypothesis can be demonstrated to be qualitatively incorrect and an entirely new event in the Kashmir region was built from a single arrival in this list (the first P-arrival at WRA).

Approved for public release; distribution is unlimited.

The analyst-reviewed location estimates (Fig. 20) indicate that the Kashmir event has been reconstructed manually entirely from the Warramunga P-detection (no other detections in the SL3 event hypothesis are now associated with the REB event). The overlapping waveforms are displayed for a number of stations in Fig. 21. It is demonstrated in Fig. 22 how a fully automatic event location is possible by taking the WRA P-detection as a trigger and projecting this back over a source region and creating multiple event hypotheses and analyzing which event hypotheses can associate the greatest number of phases and the smallest time residuals.

This is a pathological example of completely unrelated overlapping events which are likely to provide a challenge to a context-based detection framework. These events will appear very closely in slowness space for almost all global arrays and may expose limiting factors for a number of the techniques developed here. Additional examples of this nature will be sought in order to evaluate strategies for dealing with such cases.

Approved for public release; distribution is unlimited.

EVENT 3441778 HINDU KUSH REGION, AFGHANISTAN

Date	Time	Err	RMS	Latitude	Longitude	Smaj	Smin	Az	Depth	Err	Ndef	Nsta	Gap	mdist	Mdist	Qual	Author	OrigID
2005/11/05	23:24:23.92	4.21	0.89	36.5181	70.8303	23.4	13.6	2	246.1	39.2	18	17	128	13.36	149.70	m i uk	IDC_REB:I	3443754

Magnitude	Err	Nsta	Author	OrigID
mb 3.5	0.1	12	IDC_REB:I	3443754
mb1 3.6	0.1	16	IDC_REB:I	3443754
mb1mx 3.5	0.1	23	IDC_REB:I	3443754
mbtmp 4.2	0.1	16	IDC_REB:I	3443754

Sta	Dist	EvAz	Phase	Time	TRes	Azim	AzRes	Slow	SRes	Def	SNR	Amp	Per	Qual	Magnitude	ArrID
MKAR	13.36	36.2	P	23:27:22.650	-1.6	223.4	-0.5	17.1	4.8	T__	53.9	0.9	0.33	a__	mb1 3.4	26110551
BVAR	16.51	359.1	P	23:27:59.762	-0.5	158.8	-20.0	12.4	1.4	T__	49.8	2.6	0.33	a__	mb1 3.8	26112040
BVAR	16.51	359.1	S	23:30:59.262	-1.5	172.4	-6.3	21.8	1.7	T__	5.0			__		26147621
AKTO	16.70	330.4	P	23:28:02.375	-0.0	137.6	-4.0	10.9	-0.0	T__	33.7	1.6	0.33	a__	mb1 3.9	26111681
ZAL	19.94	24.7	P	23:28:36.850	-0.4	248.8	-168.	6.0	-4.8	T__	19.7	2.7	0.33	a__	mb1 3.9	26110609
ZAL	19.94	24.7	S	23:32:04.175	-4.3	159.1	-78.9	21.1	2.4	___	7.5	0.3	0.33	a__		26110694
SONM	28.44	55.4	P	23:29:57.560	1.5	270.3	10.6	9.1	0.3	T__	10.3	1.0	0.59	a__	mb 3.3	26110501
BRTR	29.30	287.7	P	23:30:04.550	0.8	100.8	16.5	8.2	-0.6	T__	8.5	1.1	0.63	a__	mb 3.3	26110565
AKASG	32.83	308.8	P	23:30:33.900	-0.3	80.6	-18.8	7.8	-0.8	T__	6.7	0.4	0.35	a__	mb 2.9	26110563
MLR	34.65	299.2	P	23:30:51.834	1.7	169.4	80.6	3.3	-5.2	T__	4.9	2.4	0.88	a__	mb 3.6	26111758
FINES	37.38	326.1	P	23:31:12.700	-0.1	114.6	-4.0	10.6	0.0	T__	12.0	2.0	0.67	a__	mb 3.5	26110539
ARCES	41.07	337.6	P	23:31:43.250	0.1	111.7	-1.8	7.5	-0.1	T__	29.2	2.6	0.44	a__	mb 3.8	26110680
HFS	42.97	321.9	P	23:31:57.994	-0.5	97.3	0.4	10.2	2.2	T__	15.5	1.8	0.37	a__	mb 3.7	26111908
NOA	44.28	323.0	P	23:32:08.250	-0.7	96.5	-0.0	7.7	-0.0	T__	7.6	1.6	0.61	a__	mb 3.5	26110692
INK	73.92	9.2	P	23:35:32.000	0.4	328.7	-11.0	4.9	-0.9	T__	11.8	1.8	0.61	a__	mb 3.7	26112063
DBIC	74.56	266.9	P	23:35:35.250	-1.0	213.3	178.2	18.0	12.3	T__	4.2	2.7	0.74	a__	mb 3.8	26110655
YKA	81.26	2.6	P	23:36:11.945	-0.1	348.9	-3.3	5.4	0.0	T__	6.0	0.6	0.34	a__	mb 3.4	26110683
WRA	82.12	121.8	P	23:36:16.409	-0.9	323.2	1.0	4.3	-0.4	T__	3.4	0.5	0.63	__	mb 3.2	26147622
PLCA	149.70	250.1	PKPbc	23:43:44.400	0.4	152.9	67.5	8.2	5.7	T__	4.5	3.1	0.91	a__		26110706

EVENT 3441926 PAKISTAN

Date	Time	Err	RMS	Latitude	Longitude	Smaj	Smin	Az	Depth	Err	Ndef	Nsta	Gap	mdist	Mdist	Qual	Author	OrigID
2005/11/05	23:25:36.11	1.52	0.56	34.2471	73.5121	36.3	24.2	43	0.0f		9	9	200	14.19	79.05	m i uk	IDC_REB:I	3443757

Magnitude	Err	Nsta	Author	OrigID
ML 4.2	0.0	2	IDC_REB:I	3443757
mb 4.0	0.1	6	IDC_REB:I	3443757
mb1 4.1	0.1	9	IDC_REB:I	3443757
mb1mx 3.9	0.1	21	IDC_REB:I	3443757
mbtmp 4.0	0.1	9	IDC_REB:I	3443757

Sta	Dist	EvAz	Phase	Time	TRes	Azim	AzRes	Slow	SRes	Def	SNR	Amp	Per	Qual	Magnitude	ArrID
MKAR	14.19	25.3	Pn	23:28:58.999	-0.2	229.4	18.3	13.3	-0.3	T__	2.3			__	mb1 3.5	26147769
BVAR	18.90	354.2	P	23:29:58.265	0.2	164.1	-7.9	10.5	-0.7	T__	5.0	0.7	0.33	__	ML 4.2	26147768
AKTO	19.75	329.6	P	23:30:08.525	1.0	145.2	6.1	11.1	0.2	T__	5.8	0.7	0.33	a__	ML 4.2	26111771
ZAL	21.24	18.6	P	23:30:22.550	-0.9	255.7	-167.	6.3	-4.5	T__	6.2	5.5	0.37	a__	mb 4.3	26110613
SONM	28.04	51.1	P	23:31:29.580	0.5	258.2	5.1	9.4	0.5	T__	4.7	0.8	0.79	a__	mb 3.7	26110622
ARCES	44.00	337.9	P	23:33:45.067	0.3	110.6	-1.2	7.7	0.2	T__	3.7	4.5	1.02	__	mb 4.3	26147766
HFS	46.10	323.1	P	23:34:01.582	0.0	99.3	2.9	9.9	2.0	T__	9.5	1.0	0.34	a__	mb 4.1	26111909
NOA	47.41	324.2	P	23:34:11.172	-0.7	97.4	1.4	7.7	0.1	T__	2.9	0.8	0.55	__	mb 3.8	26147767
WRA	79.05	123.2	P	23:37:42.409	-0.0	320.4	-1.0	5.2	0.2	T__	6.6	1.3	0.79	a__	mb 4.1	26110688
															mb1 4.5	
															mbtmp 4.1	

Fig. 20. Extract from Reviewed Event Bulletin (REB) from the International Data Center (IDC).

The first of these events was well predicted automatically (see the SL3 page displayed in Fig. 19), and the second of these events was built manually from a single phase extracted from a false event hypothesis. The relocated events are separated by approximately 350 km, although the first event is in addition over 250 km deep.

Approved for public release; distribution is unlimited.

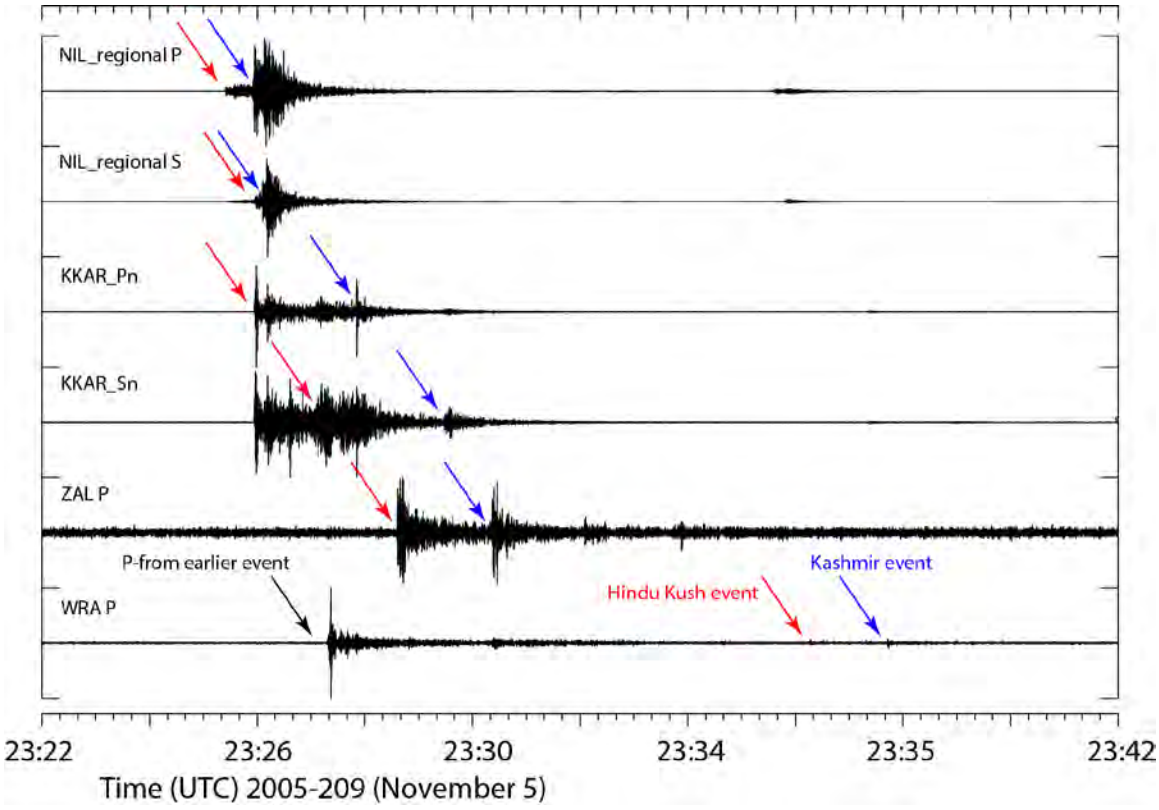

Fig. 21 Waveforms for selected stations (optimized for given phases from the Kashmir region) with arrivals from the events listed in Fig. 20 labeled.

Approved for public release; distribution is unlimited.

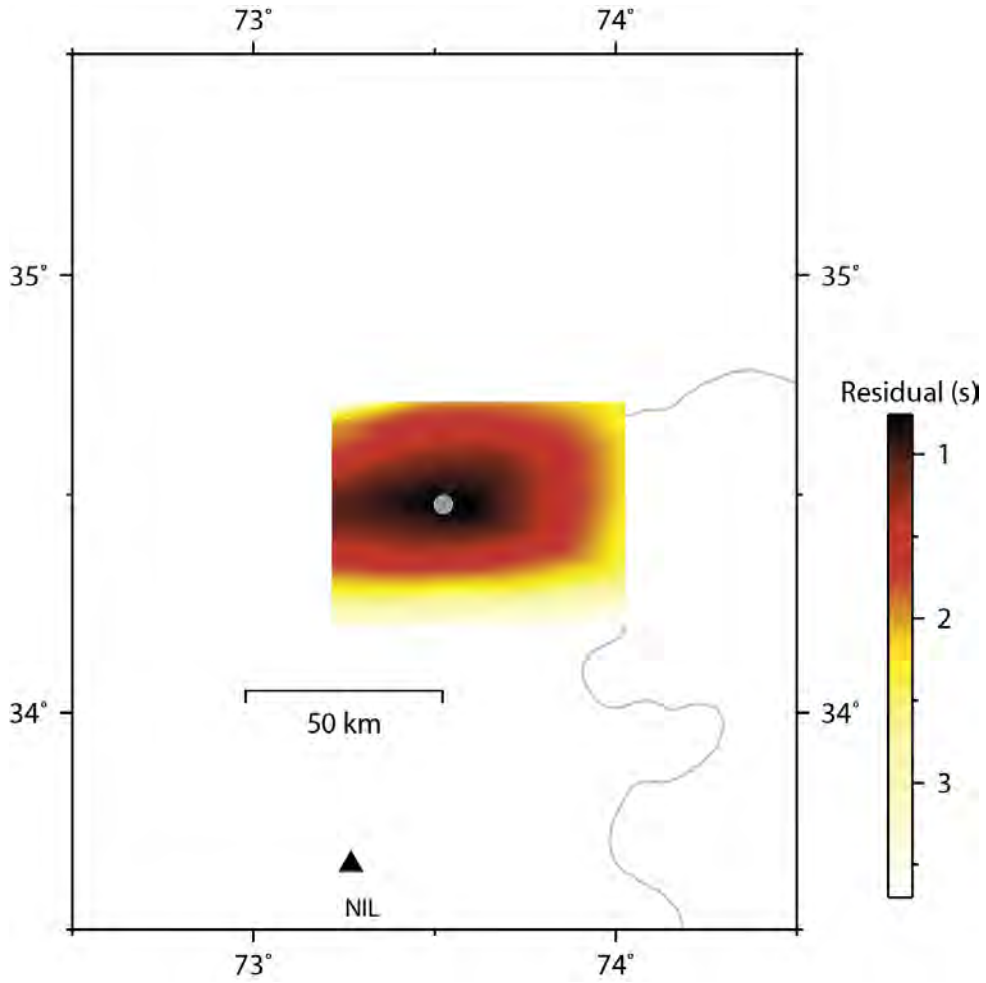

Fig. 22 Source-scan for triggering WRA P-phase at time 2005-309:23.37.43. *A geographical grid was searched with origin times determined by the theoretical traveltime to WRA and the number of automatic phase detections from other stations consistent with each of these event hypotheses was counted. Of all of the event hypotheses with the maximum number of phases (8), the hypocenter with the minimum traveltime residual was 2005-309:23.25.40 at 34.46°N, 73.60°E and 20 km deep. The colors in this plot indicate the 1-norm traveltime residual for this set of phases over a far denser grid of trial hypocenters. The optimal epicenter is at 34.477°N 73.524°E.*

.

Approved for public release; distribution is unlimited.

Lat: 36.84 Lon: 70.29 Depth: 0.0

Station	Rank	Thresh	Delta
AAK	1	2.5372	6.6351
ZALV	2	3.0180	19.8507
MKAR	3	3.0956	13.3631
AKTO	4	3.2210	16.2038
GEYT	5	3.2620	9.7500
BVAR	6	3.2949	16.1844
TORD	7	3.3250	65.1570
KURK	8	3.3777	14.9949
ARCES	9	3.3890	40.6137
HFS	10	3.4140	42.4475
FINES	11	3.4940	36.8837
SONM	12	3.5605	28.6153
YKA	13	3.5760	80.9319
ASAR	14	3.5850	84.9160
WRA	15	3.5850	82.6543
NOA	16	3.5870	43.7658
KBZ	17	3.6351	21.9364
WSAR	18	3.6359	16.8795
BRTR	19	3.6548	28.7881
ILAR	20	3.6850	74.6078

Lat: 42.698 Lon: 79.755 Depth: 0.(

Station	Rank	Thresh	Delta
AAK	1	1.3380	3.8800
MKAR	2	1.8090	4.4756
KURK	3	2.5942	7.9719
SONM	4	3.0640	19.3945
ZALV	5	3.0730	11.7473
BVAR	6	3.0928	12.0837
TORD	7	3.1760	72.4657
ARU	8	3.3590	19.3246
AKTO	9	3.3700	16.7915
ARCES	10	3.4280	38.3734
YKA	11	3.4290	74.5251
FINES	12	3.4660	36.6429
BRTR	13	3.4780	34.4904
ILAR	14	3.5000	66.8254
WRA	15	3.5570	80.0477
ASAR	16	3.5780	82.7496
NOA	17	3.6160	43.8214
GERES	18	3.6210	45.1779
HFS	19	3.6570	42.7213
CMAR	20	3.7122	29.1315

Fig. 23 Anticipated detection thresholds for the most sensitive IMS stations for the two SEL3 event hypotheses displayed in Fig. 19.

A module is being developed that, for any given event hypothesis, evaluates the phases most likely to be observed (see Fig. 23) and compares with those that were actually detected and associated. The source location for which the detection thresholds are displayed in the left hand panel of Fig. 23 is close to the estimated hypocenter of the first event hypothesis displayed in Fig. 19 and it is clear that there is a good correspondence between the observed phases and the stations with the lowest detection thresholds. The source location for which the detection thresholds are displayed in the right hand panel of Fig. 23 is close to the estimated hypocenter of the second event hypothesis displayed in Fig. 19 and it is clear that many stations which would be expected to detect an event of this magnitude are not associated with this source hypothesis. (It should be noted however that the station lists in Fig. 23 are for the current IMS and a number of these stations were not in operation at the time of the events in November 2005.)

Approved for public release; distribution is unlimited.

5 CONCLUSIONS

In the first 12 months of this contract we have made significant progress on three important components of a context-based processing pipeline: cancellation of transient signals, empirical matched field processing and adaptive beamforming, and the evaluation of event hypotheses. In the coming months, a significant modification of the framework will take place, moving from a shared-memory architecture to a message-based architecture. However, significant testing of the individual components of the pipeline will continue offline in the existing framework.

Approved for public release; distribution is unlimited.

This page is intentionally left blank.

Approved for public release; distribution is unlimited.

REFERENCES

Capon, J., Greenfield, R. J., and Kolker, R. J. (1967), Multidimensional maximum-likelihood processing of a large aperture seismic array, Proc. IEEE, 55 (2), pp. 192-211.

Harris, D. B. (1991), A waveform correlation method for identifying quarry explosions, Bull. Seismol. Society of America, Vol. 81 (6), pp. 2395-2418.

Kværna, T. and Ringdal, F. (2013), Detection Capability of the Seismic Network of the International Monitoring System for the Comprehensive Nuclear-Test-Ban Treaty, Bull. Seism. Soc. Am., Vol. 103, (2A), pp. 759-772, doi:10.1785/0120120248.

Murphy, J. R. and Barker, B. W. (2003), Revised Distance and Depth Corrections for Use in the Estimation of Short-Period P-Wave Magnitudes, Bull. Seism. Soc. Am., Vol. 93, (4), pp. 1746-1764, doi:10.1785/0120020084.

Approved for public release; distribution is unlimited.

LIST OF SYMBOLS, ABBREVIATIONS, AND ACRONYMS

CTBTO	Comprehensive Nuclear-Test-Ban Treaty Organization
IDC	International Data Center
MVDR	Minimum Variance Distortionless Receiver
REB	Reviewed Event Bulletin (an analyst reviewed event bulletin generated at the IDC)
SEL3	Standard Event List 3 (a fully automatic seismic event bulletin generated at the IDC)

Approved for public release; distribution is unlimited.

DISTRIBUTION LIST

DTIC/OCP
8725 John J. Kingman Rd, Suite 0944
Ft Belvoir, VA 22060-6218 1 cy

AFRL/RVIL
Kirtland AFB, NM 87117-5776 2 cys

Official Record Copy
AFRL/RVBYE/Robert Raistrick 1 cy

Approved for public release; distribution is unlimited.

This page is intentionally left blank.

Approved for public release; distribution is unlimited.